



(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *with international search report*

1 HIERARCHICAL DATA-DRIVEN SEARCH AND NAVIGATION SYSTEM AND
2 METHOD FOR INFORMATION RETRIEVAL

3 This application is a continuation-in-part of App. Ser. No. 09/961,131, entitled "Scalable
4 Hierarchical Navigation System and Method for Information Retrieval," filed September
5 21, 2001, which is a continuation-in-part of Application Ser. No. 09/573,305, entitled
6 "Hierarchical Data-Driven Navigation System and Method for Information Retrieval,"
7 filed May 18, 2000, which are incorporated herein by this reference.

8 1. Field of the Invention

9 The present invention generally relates to information search and navigation
10 systems.

11 2. Background of the Invention

12 Information retrieval from a database of information is an increasingly
13 challenging problem, particularly on the World Wide Web (WWW), as increased
14 computing power and networking infrastructure allow the aggregation of large amounts
15 of information and widespread access to that information. A goal of the information
16 retrieval process is to allow the identification of materials of interest to users.

17 As the number of materials that users may search and navigate increases,
18 identifying relevant materials becomes increasingly important, but also increasingly
19 difficult. Challenges posed by the information retrieval process include providing an
20 intuitive, flexible user interface and completely and accurately identifying materials
21 relevant to the user's needs within a reasonable amount of time. Another challenge is to
22 provide an implementation of this user interface that is highly scalable, so that it can

1 readily be applied to the increasing amounts of information and demands to access that
2 information. The information retrieval process comprehends two interrelated technical
3 aspects, namely, information organization and access.

4 Current information search and navigation systems usually follow one of three
5 paradigms. One type of information search and navigation system employs a database
6 query system. In a typical database query system, a user formulates a structured query by
7 specifying values for fixed data fields, and the system enumerates the documents whose
8 data fields contain those values. PriceSCAN.com uses such an interface, for example.
9 Generally, a database query system presents users with a form-based interface, converts
10 the form input into a query in a formal database language, such as SQL, and then executes
11 the query on a relational database management system. Disadvantages of typical query-
12 based systems include that they allow users to make queries that return no documents and
13 that they offer query modification options that lead only to further restriction of the result
14 set (the documents that correspond to the user's specifications), rather than to expansion
15 or extension of the result set. In addition, database query systems typically exhibit poor
16 performance for large data sets or heavy access loads; they are often optimized for
17 processing transactions rather than queries.

18 A second type of information search and navigation system is a free-text search
19 engine. In a typical free-text search engine, the user enters an arbitrary text string, often
20 in the form of a Boolean expression, and the system responds by enumerating the
21 documents that contain matching text. Google.com, for example, includes a free-text
22 search engine. Generally a free-text search engine presents users with a search form,

1 often a single line, and processes queries using a precomputed index. Generally this
2 index associates each document with a large portion of the words contained in that
3 document, without substantive consideration of the document's content. Accordingly, the
4 result set is often a voluminous, disorganized list that mixes relevant and irrelevant
5 documents. Although variations have been developed that attempt to determine the
6 objective of the user's query and to provide relevance rankings to the result set or to
7 otherwise narrow or organize the result set, these systems are limited and unreliable in
8 achieving these objectives.

9 A third type of information search and navigation system is a tree-based directory.
10 In a tree-based directory, the user generally starts at the root node of the tree and specifies
11 a query by successively selecting refining branches that lead to other nodes in the tree.
12 Shopping.yahoo.com uses a tree-based directory, for example. In a typical
13 implementation, the hard-coded tree is stored in a data structure, and the same or another
14 data structure maps documents to the node or nodes of the tree where they are located. A
15 particular document is typically accessible from only one or, at most, a few, paths through
16 the tree. The collection of navigation states is relatively static—while documents are
17 commonly added to nodes in the directory, the structure of the directory typically remains
18 the same. In a pure tree-based directory, the directory nodes are arranged such that there
19 is a single root node from which all users start, and every other directory node can only be
20 reached via a unique sequence of branches that the user selects from the root node. Such
21 a directory imposes the limitation that the branches of the tree must be navigationally
22 disjoint—even though the way that documents are assigned to the disjoint branches may

1 not be intuitive to users. It is possible to address this rigidity by adding additional links to
2 convert the tree to a directed acyclic graph. Updating the directory structure remains a
3 difficult task, and leaf nodes are especially prone to end up with large numbers of
4 corresponding documents.

5 In all of these types of search and navigation systems, it may be difficult for a user
6 to revise a query effectively after viewing its result set. In a database query system, users
7 can add or remove terms from the query, but it is generally difficult for users to avoid
8 underspecified queries (i.e. too many results) or overspecified queries (i.e. no results).
9 The same problem arises in free-text search engines. In tree-based directories, the only
10 means for users to revise a query is either to narrow it by selecting a branch or to
11 generalize it by backing up to a previous branch.

12 Having an effective means of revising queries is useful in part because users often
13 do not know exactly what they are looking for. Even users who do know what they are
14 looking for may not be able to express their criteria precisely. And the state of the art in
15 information retrieval technology cannot guarantee that even a precisely stated query will
16 be interpreted as intended by the user. Indeed, it is unlikely that a perfect means for
17 formation of a query even exists in theory. As a result, it is helpful that the information
18 retrieval process be a dialogue with interactive responses between the user and the
19 information retrieval system. This dialogue model may be more effectively implemented
20 with an effective query revision process.

21 Some information retrieval systems combine a search engine with a vocabulary of
22 words or phrases used to classify documents. These systems enable a three-step process

1 for information retrieval. In the first step, a user enters a text query into a search form, to
2 which the system responds with a list of matching vocabulary terms. In the second step,
3 the user selects from this list, to which the system responds with a list of documents.
4 Finally, in the third step, the user selects a document.

5 A problem with such systems is that they typically do not consider the possibility
6 that a user's search query may match a conjunction of two or more vocabulary terms,
7 rather than an individual term. For example, in a system whose vocabulary consists of
8 consumer electronics products and manufacturers, a search for *Sony DVD players*
9 corresponds to a conjunction of two vocabulary terms: *Sony* and *DVD players*. Some
10 systems may address this problem by expanding their vocabularies to include vocabulary
11 terms that incorporate compound concepts (e.g., all valid combinations of manufacturers
12 and products), but such an exhaustive approach is not practical when there are a large
13 number of independent concepts in a system, such as product type, manufacturer, price,
14 condition, etc. Such systems also may fail to return concise, usable search results,
15 partially because the number of compound concepts becomes unmanageable. For
16 example, a search for *software* in the Yahoo category directory returns 477 results, most
17 of which represent compound concepts (e.g., *Health Care > Software*).

18 Various other systems for information retrieval are also available. For example,
19 U.S. Patents Nos. 5,715,444 and 5,983,219 to Danish et al., both entitled "Method and
20 System for Executing a Guided Parametric Search," disclose an interface for identifying a
21 single item from a family of items. The interface provides users with a set of lists of
22 features present in the family of items and identifies items that satisfy selected features.

1 Other search and navigation systems include i411's Discovery Engine, Cybrant's
2 Information Engine, Mercado's IntuiFind, and Requisite Technology's BugsEye.

3 3. Summary of the Invention

4 The present invention, a highly scalable, hierarchical, data-driven information
5 search and navigation system and method, enables the search and navigation of a
6 collection of documents or other materials using certain common attributes associated
7 with those materials. The search interface allows the user to enter queries that may
8 correspond to either single terms or combinations of terms from a vocabulary used to
9 classify documents. The navigation interface allows the user to select values for the
10 attributes associated with the materials in the current navigation state and returns the
11 materials that correspond to the user's selections. In some embodiments, the user's
12 selections may be combined using Boolean operators. The present invention enables this
13 navigation mode by associating terms (attribute-value pairs) with the documents, defining
14 a set of hierarchical refinement relationships (i.e., a partial order) among the terms, and
15 providing a guided navigation mechanism based on the association of terms with
16 documents and the relationships among the terms.

17 The present invention includes several components and features relating to a
18 hierarchical data-driven search and navigation system. Among these are a user interface,
19 a knowledge base, a process for generating and maintaining the knowledge base, a
20 navigable data structure and method for generating the data structure, WWW-based
21 applications of the system, and methods of implementing the system. Although the
22 invention is described herein primarily with reference to a WWW-based system for

1 navigating a product database, it should be understood that a similar search and
2 navigation system could be employed in any database context where materials may be
3 associated with terms and users can identify materials of interest by way of those terms.

4 The present invention uses a knowledge base of information regarding the
5 collection of materials to formulate and to adapt the interface to guide the user through
6 the collection of navigation states by providing relevant navigation options. The
7 knowledge base includes an enumeration of attributes relevant to the materials, a range of
8 values for each attribute, and a representation of the partial order that relates terms (the
9 attribute-value pairs). Attribute-value pairs for materials relating to entertainment, for
10 example, may be *Products: Movies* and *Director: Spike Lee*. (Attribute-value pairs are
11 represented throughout this specification in this *Attribute: Value* format; navigation
12 states are represented as bracketed expressions of attribute-value pairs.) The knowledge
13 base also includes a classification mapping that associates each item in the collection of
14 materials with a set of terms that characterize that item.

15 The knowledge base is typically organized by domains, which are sets of
16 materials that conform to natural groupings. Preferably, a domain is chosen such that a
17 manageable number of attributes suffice to effectively distinguish and to navigate among
18 the materials in that domain. The knowledge base preferably includes a characterization
19 of each domain, which might include rules or default expectations concerning the
20 classification of documents in that domain. A particular item may be in more than one
21 domain.

1 Embodiments of the present invention include a user interface for searching. This
2 interface allows users to use a free-text search to find terms of interest. A free-text query
3 may be composed of one or more words. The system may interpret free-text queries in
4 various ways; the interpretations used to execute a free-text query will determine the
5 nature of the search results for that query. A single-term interpretation maps the complete
6 query to an individual term in the knowledge base. A multi-term interpretation maps the
7 query to a conjunction of two or more terms in the knowledge base—that is, a plurality of
8 terms that corresponds to a conjunctive navigation state. Depending on the particular
9 implementation and application context, a free-text query may be mapped to one or more
10 single-term interpretations, one or more multi-term interpretations, or a combination of
11 both. In another aspect of the present invention, the user interface allows users to use a
12 free-text search either merely to find matching terms or further to find navigation states or
13 materials associated with the matching terms.

14 The present invention also includes a user interface for navigation. The user
15 interface preferably presents the user's navigation state as an expression of terms
16 organized by attribute. For a given expression of terms, the user interface presents
17 materials that are associated with those terms in accordance with that expression and
18 presents relevant navigation options for narrowing or for generalizing the navigation
19 state. In one aspect of the present invention, users navigate through the collection of
20 materials by selecting and deselecting terms.

21 In one aspect of the present invention, the user interface responds immediately to
22 the selection or the deselection of terms, rather than waiting for the user to construct and

1 to submit a comprehensive query composed of multiple terms. Once a query has been
2 executed, the user may narrow the navigation state by conjunctively selecting additional
3 terms, or by refining existing terms. Alternatively, the user may broaden the navigation
4 state by deselecting terms that have already been conjunctively selected or by
5 generalizing the terms. In preferred embodiments, the user may broaden the navigation
6 state by deselecting terms in an order different from that in which they were conjunctively
7 selected. For example, a user could start at {*Products: Movies*}, narrow by conjunctively
8 selecting an additional term to {*Products: Movies AND Genre: Drama*}, narrow again to
9 {*Products: Movies AND Genre: Drama AND Director: Spike Lee*}, and then broaden by
10 deselecting a term to {*Products: Movies AND Director: Spike Lee*}.

11 In another aspect of the present invention, the user may broaden the navigation
12 state by disjunctively selecting additional terms. For example, a user could start at
13 {*Products: DVDs*}, and then broaden by disjunctively selecting a term to {*Products:*
14 *DVDs OR Products: Videos*}, and then narrow by conjunctively selecting a term to
15 {(*Products: DVDs OR Products: Videos*) AND *Director: Spike Lee*}.

16 In another aspect of the present invention, the user may narrow the navigation
17 state by negationally selecting additional terms. For example, a user could start at
18 {*Products: DVDs*}, narrow by conjunctively selecting a term to {*Products: DVDs AND*
19 *Genre: Comedy*}, and then narrow by negationally selecting a term to {*Products: DVDs*
20 *AND Genre: Comedy AND (NOT Director: Woody Allen)*}.

21 In another aspect of the present invention, the user interface presents users with
22 context-dependent navigation options for modifying the navigation state. The user

1 interface does not present the user with options whose selection would correspond to no
2 documents in the resulting navigation state. Also, the user interface presents new
3 navigation options as they become relevant. The knowledge base may contain rules that
4 determine when particular attributes or terms are made available to users for navigation.

5 In another aspect of the invention—for example, when the materials correspond to
6 products available for purchase from various sources—the knowledge base includes a
7 catalog of canonical representations that have been aggregated from the materials.

8 In another aspect of the invention, the knowledge base may include definitions of
9 stores, sets of materials that are grouped to be retrievable at one time. A store may
10 include documents from one or more domains. An item may be assigned to more than
11 one store. The knowledge base may also include rules to customize navigation for
12 particular stores.

13 In another aspect of the invention, the knowledge base is developed through a
14 multi-stage, iterative process. Workflow management allocates resources to maximize
15 the efficiency of generating and of maintaining the knowledge base. The knowledge base
16 is used to generate data structures that support navigation through a collection of
17 materials. In one aspect of the invention, the system includes a hierarchy (i.e., a partial
18 order) of navigation states that map expressions of terms to the sets of materials with
19 which those terms are associated. In another aspect of the invention, the navigation states
20 are related by transitions corresponding to terms used to narrow or broaden from one
21 navigation state to another. The navigation states may be fully or partially precomputed,
22 or may be entirely computed at run-time. In another aspect of the invention,

1 implementations of the invention may be scalable through parallel or distributed
2 computation. In addition, implementations of the invention may employ master and slave
3 servers arranged in a hierarchical configuration.

4 4. Brief Description of the Drawings

5 The invention, including these and other features thereof, may be more fully
6 understood from the following description and accompanying drawings, in which:

7 Figure 1 is a view of a user interface to a search and navigation system in
8 accordance with an embodiment of the present invention.

9 Figure 2 is a view of the user interface of Figure 1, showing a drop-down pick list
10 of navigable terms.

11 Figure 3 is a view of the user interface of Figure 1, showing a navigation state.

12 Figure 4 is a view of the user interface of Figure 1, showing a navigation state.

13 Figure 5 is a view of the user interface of Figure 1, showing a navigation state.

14 Figure 6 is a view of the user interface of Figure 1, showing a navigation state.

15 Figure 7 is a view of the user interface of Figure 1, showing a navigation state.

16 Figure 8 is a view of the user interface of Figure 1, showing a navigation state.

17 Figure 9 is a view of the user interface of Figure 1, showing the result of a free-
18 text search.

19 Figure 10 is a view of a user interface in accordance with another embodiment of
20 the present invention, showing the result of a free-text search.

21 Figure 11 is a view of the user interface of Figure 10, showing the result of a free-
22 text search.

1 Figure 12 is a view of a user interface in accordance with another embodiment of
2 the invention, showing the result of a free-text search.

3 Figure 13 is a view of a user interface in accordance with another embodiment of
4 the invention, showing information about a particular document.

5 Figures 14A-C are representative examples of how the range of values for an
6 attribute could be partially ordered in accordance with an embodiment of the present
7 invention.

8 Figure 15 is a block diagram of a process for collecting and classifying documents
9 in accordance with an embodiment of the present invention.

10 Figure 16 is a table illustrating how a set of documents may be classified in
11 accordance with an embodiment of the present invention.

12 Figure 17 is a representative partial order of navigation states in accordance with
13 an embodiment of the present invention.

14 Figure 18 is a block diagram of a process for precomputing a navigation state in
15 accordance with an embodiment of the present invention.

16 Figure 19 is a view of a user interface to a search and navigation system in
17 accordance with an embodiment of the invention, showing disjunctive selection.

18 Figure 20 is a view of a user interface to a search and navigation system in
19 accordance with an embodiment of the invention, showing disjunctive selection.

20 Figure 21 is a view of a user interface to a search and navigation system in
21 accordance with an embodiment of the invention, showing negational selection.

1 Figure 22 is a view of a user interface to a search and navigation system in
2 accordance with an embodiment of the invention, showing negational selection.

3 Figure 23 is a block diagram of a method for processing a free-text search query
4 in accordance with an embodiment of the present invention.

5 Figure 24 is a block diagram of a system and a method for processing a request
6 across multiple servers in accordance with an embodiment of the present invention.

7 Figure 25 is a flow diagram of steps for combining refinement options from slave
8 servers in accordance with an embodiment of the present invention.

9 5. Detailed Description of the Preferred Embodiments

10 User Interface

11 In accordance with one embodiment of the present invention, Figure 1 shows a
12 user interface 10 to a hierarchical, data-driven search and navigation system. The search
13 and navigation system operates on a collection of documents defined in a knowledge
14 base. As is shown, the user is preferably presented with at least two alternative methods
15 of using the search and navigation system: (1) by selecting terms to navigate through the
16 collection of documents, or (2) by entering a desired query in a search box.

17 The search and navigation system preferably organizes documents by domain. In
18 accordance with one embodiment of the present invention, the user interface 10 shown in
19 Figures 1-9 is operating on a set of documents that are part of a wine domain. Preferably,
20 a domain defines a portion of the collection of documents that reflects a natural grouping.
21 Generally, the set of attributes used to classify documents in a domain will be a
22 manageable subset of the attributes used to classify the entire collection of documents. A

1 domain definition may be a type of product, e.g., wines or consumer electronics. A
2 domain may be divided into subdomains to further organize the collection of documents.
3 For example, there can be a consumer electronics domain that is divided into the
4 subdomains of televisions, stereo equipment, etc. Documents may correspond to goods
5 or services.

6 The user interface may allow users to navigate in one domain at a time.

7 Alternatively, the user interface may allow the simultaneous navigation of multiple
8 domains, particularly when certain attributes are common to multiple domains.

9 The user interface allows the user to navigate through a collection of navigation
10 states. Each state is composed of an expression of terms and of the set of documents
11 associated with those terms in accordance with that expression. In the embodiment
12 shown in Figures 1-9, users navigate through the collection of navigation states by
13 conjunctively selecting and deselecting terms to obtain the navigation state corresponding
14 to each expression of conjunctively selected terms. Preferably, as in Figure 4, the user
15 interface 10 presents a navigation state by displaying both the list 50 of terms 52 and a list
16 41 of some or all of the documents 42 that correspond to that state. Preferably, the user
17 interface presents the terms 52 of the navigation state organized by attribute. Preferably,
18 the initial navigation state is a root state that corresponds to no term selections and,
19 therefore, to all of the documents in the collection.

20 As shown in Figure 2, the user interface 10 allows users to narrow the navigation
21 state by choosing a value 28 for an attribute 22, or by replacing the currently selected
22 value with a more specific one (if appropriate). Preferably, the user interface 10 presents

1 users with the options available to narrow the present navigation state, preferably with
2 relevant terms organized by attribute. In some embodiments of the present invention, as
3 shown in Figure 2, users can select values 28 from drop-down lists 26 denoted by
4 indicators 24, that are organized by attributes 22 in the current navigation state. The user
5 interface may present these navigation options in a variety of formats. For example,
6 values can be presented as pictures or as symbols rather than as text. The interface may
7 allow for any method of selecting terms, e.g., mouse clicks, keyboard strokes, or voice
8 commands. The interface may be provided through various media and devices, such as
9 television or WWW, and telephonic or wireless devices. Although discussed herein
10 primarily as a visual interface, the interface may also include an audio component or be
11 primarily audio-based.

12 Preferably, in the present navigation state, the user interface only presents options
13 for narrowing the navigation state that lead to a navigation state with at least one
14 document. This preferred criteria for providing navigation options ensures that there are
15 no "dead ends," or navigation states that correspond to an empty result set.

16 Preferably, the user interface only presents options for narrowing the navigation
17 state if they lead to a navigation state with strictly fewer documents than the present one.
18 Doing so ensures that the user interface does not present the user with choices that are
19 already implied by terms in the current navigation state.

20 Preferably, the user interface presents a new navigation state as soon as the user
21 has chosen a term 28 to narrow the current navigation state, without any further triggering

1 action by the user. Because the system responds to each user with immediate feedback,
2 the user need not formulate a comprehensive query and then submit the query.

3 In accordance with one embodiment of the present invention, as shown in
4 Figures 3 and 4, the user interface 10 may enable broadening of the current navigation
5 state by allowing the user to remove terms 52 from the list 50 of terms conjunctively
6 selected. For example, the interface 10 may provide a list 50 with checkboxes 54 for
7 removing selections and a button 56 to trigger the computation of the new navigation
8 state. In the illustrated embodiment, the user can remove conjunctively selected terms 52
9 in any order and can remove more than one selection 52 at a time.

10 Preferably, the navigation options presented to the user are context-dependent.
11 For example, terms that refine previously selected terms may become navigation options
12 in the resulting navigation state. For example, referring to Figure 5, after the term
13 *Flavors: Wood and Nut Flavors* 52 is conjunctively selected (the user has selected the
14 value Wood and Nut Flavors 23 for the attribute Flavors), Wood and Nut Flavors 23 then
15 appears in the interface for the new navigation state in the list 20 of attributes and allows
16 conjunctive selection of values 28 that relate to that specific attribute for further
17 refinement of the query. The user interface may also present certain attributes that were
18 not presented initially, as they become newly relevant. For example, comparing Figure 3
19 to Figure 2, the attribute French Vineyards 25 appears in the list 20 of attributes only after
20 the user has already conjunctively selected the term *Regions: French Regions* in a
21 previous navigation state. Attributes may be embedded in this way to as many levels as
22 are desired. Presenting attributes as navigation options when those attributes become

1 relevant avoids overwhelming the user with navigation options before those options are
2 meaningful.

3 Additionally, for some attributes 22, multiple incomparable (non-refining)
4 conjunctive selections of values 28 may be applicable. For example, for the attribute
5 Flavor, the values Fruity and Nutty, neither of which refines the other, may both be
6 conjunctively selected so that the terms *Flavors: Fruity* and *Flavors: Nutty* narrow the
7 navigation state. Thus, users may sometimes be able to refine a query by conjunctively
8 selecting multiple values under a single attribute.

9 Preferably, certain attributes will be eliminated as navigation options if they are
10 no longer valid or helpful choices. For example, if all of the documents in the result set
11 share a common term (in addition to the term(s) selected to reach the navigation state),
12 then conjunctive selection of that term will not further refine the result set; thus, the
13 attribute associated with that term is eliminated as a navigation option. For example,
14 comparing Figure 6 with Figure 4, the attribute Wine Types 27 has been eliminated as a
15 navigation option because all of the documents 42 in the result set share the same term,
16 *Wine Types: Appellational Wines*. In preferred embodiments, an additional feature of the
17 interface 10 is that this information is presented to the user as a common characteristic of
18 the documents 42 in the result set. For example, referring to Figure 6, the interface 10
19 includes a display 60 that indicates the common characteristics of the documents 42 in the
20 result set. Removing a term as a navigation option when all of the documents in the result
21 set share that term prevents the user from wasting time by conjunctively selecting terms
22 that do not refine the result set.

1 Preferably, the user interface also eliminates values as navigation options if their
2 selection would result in no documents in the result set. For example, comparing Figure
3 8 to Figure 7, after the user selects the term *Wine Spectator Range: 95 - 100*, the user
4 interface eliminates as navigation options all the values 28, 29 in the list 26 of values for
5 the attribute Appellations 22 except for the values Alexander Valley 29 and Napa Valley
6 29. Alexander Valley 29 and Napa Valley 29 are the only two values in the list 26 of
7 values for the attribute Appellations that return at least one document in the result set; all
8 other values 28 return the empty set. Removing values as navigation options that would
9 result in an empty result set saves the user time by preventing the user from reaching
10 dead-ends.

11 Preferably, the user interface allows users to enter free-text search queries that
12 may be composed of one or more words. The system may interpret free text queries in
13 various ways. In particular, the system may map a free-text query to two types of search
14 results: single-term interpretations and multi-term interpretations. A single-term
15 interpretation maps the complete query to an individual term in the knowledge base. A
16 multi-term interpretation maps the query to a conjunction of two or more terms in the
17 knowledge base—that is, a plurality of terms that corresponds to a conjunctive navigation
18 state. Depending on the particular implementation and application context, a free-text
19 query may be mapped to one or more single-term interpretations, one or more multi-term
20 interpretations, or a combination of both types of interpretations. In another aspect of the
21 present invention, the user interface allows users to use a free-text search either to find
22 matching terms or further to find materials associated with matching terms.

1 In accordance with one embodiment of the present invention, illustrated in Figure
2 9, in interface 90, a search box 30 preferably allows users to perform a free-text search for
3 terms of interest, rather than performing a full-text search of the documents themselves.
4 Preferably, the user interface responds to such a search by presenting a list 32 of single-
5 term interpretations 33 including terms organized by attribute 36, and allowing the user to
6 select from among them. Preferably, the user interface responds to the user's selection by
7 presenting the user with the navigation state corresponding to the selection of that term.
8 The user may then either navigate from that state (i.e., by narrowing or broadening it) or
9 perform additional free-text searches for terms.

10 In accordance with another embodiment of the present invention, illustrated in
11 Figure 10, the user interface 100 responds to free-text search queries by presenting a list
12 32 of multi-term interpretations 34, and allowing the user to select from among them.
13 Preferably, the user interface responds to the user's selection by presenting the user with
14 the navigation state corresponding to the selection of that conjunction of terms. The user
15 may then either navigate from that state (i.e., by narrowing or broadening it) or perform
16 additional free-text searches for terms.

17 In accordance with another embodiment of the present invention, illustrated in
18 Figure 11, the user interface 100 responds to free-text search queries by presenting a list
19 32 of single-term interpretations 33 and multi-term interpretations 34, and allowing the
20 user to select from among them. Preferably, the user interface responds to the user's
21 selection by presenting the user with the navigation state corresponding to the selection of

1 that term or conjunction of terms. The user may then either navigate from that state (i.e.,
2 by narrowing or broadening it) or perform additional free-text searches for terms.

3 In accordance with another embodiment of the present invention, illustrated in
4 Figure 12, the user interface 105 responds to free-text search queries by directly
5 presenting the set of matching documents 35, for example, in accordance with full-text
6 search of the documents. The user may then either navigate from that result (i.e., by
7 narrowing or broadening it) or perform additional free-text searches for terms.

8 Preferably, the user interface 10 presents a full or partial list 41 of the documents
9 that correspond to the current navigation state. Preferably, if a user is interested in a
10 particular document 42, the user may select it and obtain a record 70 containing further
11 information about it, including the list 72 of terms 74 that are associated with that
12 document, as shown in Figure 13. Preferably, the user interface 10 allows the user to
13 conjunctively select any subset of those terms 74 and thereby navigate to the navigation
14 state that corresponds to the selected term expression.

15 Preferably, the user interface 10 also offers navigation options that directly link to
16 an associated navigation state that is relevant to, but not necessarily a generalization or
17 refinement of, the present navigation state. These links preferably infer the user's
18 interests from the present navigation state and enable the user to cross-over to a related
19 topic. For example, if the user is visiting a particular navigation state in a food domain,
20 links may direct the user to navigation states of wines that would complement those foods
21 in the wine domain.

1 In accordance with another embodiment of the present invention, the user is
2 preferably presented with additional methods of using the search and navigation system
3 such as: (1) by conjunctively selecting terms, (2) by disjunctively selecting terms, (3) by
4 negationally selecting terms, or (4) by entering a desired keyword in a search box.

5 In another aspect of the present invention, the user may broaden the navigation
6 state by disjunctively selecting additional terms. For example, a user could start at
7 {*Products: DVDs*}, and then broaden by disjunctively selecting a term to {*Products:*
8 *DVDs OR Products: Videos*}, and then narrow by conjunctively selecting a term to
9 {(*Products: DVDs OR Products: Videos*) AND *Director: Spike Lee*}. Figure 19 shows a
10 user interface 300 to a hierarchical, data-driven search and navigation system. The user
11 interface 300 is operating on a collection of records relating to mutual funds. The
12 interface 300 presents navigation options, including a list of attributes 310 relating to
13 mutual funds and a list of terms 314 for a particular attribute 312, such as Fund Family,
14 under consideration by a user. A selected term 316 is highlighted. As shown, the
15 attribute-value pair {*Fund Family: Fidelity Investments*} has previously been selected.
16 The illustrated search and navigation system allows the user to select attribute-value pairs
17 disjunctively. As shown in Figure 20, after the user subsequently selects {*Fund Family:*
18 *Vanguard Group*} in addition, the interface 300 presents a new navigation state {*Fund*
19 *Family: Fidelity Investments OR Fund Family: Vanguard Group*}, including mutual
20 funds 320 that match either selected attribute-value pair. Accordingly, both selected
21 attribute-value pairs 316 are highlighted. In some embodiments, for example, to reduce
22 computational requirements, disjunctive combination of attribute-value pairs may be

1 limited to mutually incomparable attribute-value pairs that correspond to the same
2 attribute.

3 In another aspect of the present invention, the user may narrow the navigation
4 state by negationally selecting additional terms. For example, a user could start at
5 {*Products: DVDs*}, narrow by conjunctively selecting a term to {*Products: DVDs AND*
6 *Genre: Comedy*}, and then narrow by negationally selecting a term to {*Products: DVDs*
7 *AND Genre: Comedy AND (NOT Director: Woody Allen)*}. Figure 21 shows another
8 interface 400 to a hierarchical, data-driven search and navigation system. The user
9 interface 400 is operating on a collection of records relating to entertainment products.
10 The user interface 400 includes a header 410 and a navigation area 412. The header 410
11 indicates the present navigation state {*Products: DVDs AND Genre:Drama*}, and implies
12 the refinement options currently under consideration by the user. The leader "Not
13 Directed By" 414 indicates a negational operation with respect to the Director attribute.
14 The interface lists the attribute-value pairs 416 that can be combined with the expression
15 for the present navigation state under this operation. As shown in Figure 22, after the
16 user selects the term *Director: Martin Scorsese*, the interface 400 presents a new
17 navigation state {*Products: DVDs AND Genre:Drama AND (NOT Director: Martin*
18 *Scorsese)*}.

19 Although the interface to the search and navigation system has been described
20 herein as a user interface, the interface could provide other forms of access to the search
21 and navigation system. In alternative embodiments, the interface may be an applications
22 program interface to allow access to the search and navigation system for or through other

1 applications. The interface may also enhance the functionality of an independent data-
2 oriented application. The interface may also be used in the context of a WWW-based
3 application or an XML-based application. The search and navigation system may also
4 support multiple interface modes simultaneously. The search and navigation system may
5 be made available in a variety of ways, for example via wireless communications or on
6 handheld devices.

7 **Knowledge Base**

8 Preferably, the search and navigation system stores all information relevant to
9 navigation in a knowledge base. The knowledge base is the repository of information
10 from two processes: taxonomy definition and classification. Taxonomy definition is the
11 process of identifying the relevant attributes to characterize documents, determining the
12 acceptable values for those attributes (such as a list or range of values), and defining a
13 partial order of refinement relationships among terms (attribute-value pairs).
14 Classification is the process of associating terms with documents. The knowledge base
15 may also be used to maintain any information assets that support these two processes,
16 such as domains, classification rules and default expectations. Additionally, the
17 knowledge base may be used to maintain supplementary information and materials that
18 affect users' navigation experience.

19 The taxonomy definition process identifies a set of attributes that appropriately
20 characterize documents. A typical way to organize the taxonomy definition process is to
21 arrange the collections of documents into domains, which are sets of documents that
22 conform to a natural grouping and for which a manageable number of attributes suffice to

1 effectively distinguish and navigate among the documents in that domain. The knowledge
2 base preferably includes a characterization of each domain, which might include rules or
3 default expectations concerning the classification of documents in that domain.

4 The taxonomy definition process also identifies a full set of values, at varying
5 levels of specificity when appropriate, for each attribute. The values preferably identify
6 the specific properties of the documents in the collection. The values may be enumerated
7 explicitly or defined implicitly. For example, for a "color" attribute, a full set of valid
8 color values may be specified, but for a "price" or "date" attribute, a range within which
9 the values may fall or a general data type, without defining a range, may be specified.
10 The process of identifying these values may include researching the domain or analyzing
11 the collection of documents.

12 The taxonomy definition process also defines a partial order of refinement
13 relationships among terms (attribute-value pairs). For example, the term *Origin: France*
14 could refine the term *Origin: Europe*. The refinement relationship is transitive and
15 antisymmetric but not necessarily total. Transitivity means that, if term A refines term B
16 and term B refines term C, then term A refines term C. For example, if *Origin: Paris*
17 refines *Origin: France* and *Origin: France* refines *Origin: Europe*, then *Origin: Paris*
18 refines *Origin: Europe*. Antisymmetry means that, if two terms are distinct, then both
19 terms cannot refine each other. For example, if *Origin: Paris* refines *Origin: France*,
20 then *Origin: France* does not refine *Origin: Paris*.

21 Further, the partial order of refinement relationships among terms is not
22 necessarily a total one. For example, there could be two terms, *Origin: France* and

1 *Origin: Spain*, such that neither term refines the other. Two terms with this property are
2 said to be incomparable. Generally, a set of two or more terms is mutually incomparable
3 if, for every pair of distinct terms chosen from that set, the two terms are incomparable.
4 Typically, but not necessarily, two terms with distinct attributes will be incomparable.

5 Given a set of terms, a term is a maximal term in that set if it does not refine any
6 other terms in the set, and it is a minimal term in that set if no other term in the set refines
7 it. For example, in the set {*Origin: France*, *Origin: Paris*, *Origin: Spain*, *Origin:*
8 *Madrid*}, *Origin: France* and *Origin: Spain* are maximal, while *Origin: Paris* and
9 *Origin: Madrid* are minimal. In the knowledge base, a term is a root term if it does not
10 refine any other terms and a term is a leaf term if no other term refines it.

11 Figures 14A, 14B, and 14C illustrate attributes 112 and values 114, arranged in
12 accordance with the partial order relationships, that could be used for classifying wines.
13 The attributes 112 are Type/Varietal, Origin, and Vintage. Each attribute 112
14 corresponds to a maximal term for that attribute. An attribute 112 can have a flat set of
15 mutually incomparable values (e.g., Vintage), a tree of values (e.g., Origin), or a general
16 partial order that allows a value to refine a set of two or more mutually incomparable
17 values (e.g., Type/Varietal). The arrows 113 indicate the refinement relationships among
18 values 114.

19 Attributes and values may be identified and developed in several ways, including
20 manual or automatic processing and the analysis of documents. Moreover, this kind of
21 analysis may be top-down or bottom-up; that is, starting from root terms and working
22 towards leaf terms, or starting from leaf terms and working towards root terms. Retailers,

1 or others who have an interest in using the present invention to disseminate information,
2 may also define attributes and terms.

3 The classification process locates documents in the collection of navigation states
4 by associating each document with a set of terms. Each document is associated with a set
5 of mutually incomparable terms, e.g., {*Type/Varietal: Chianti, Origin: Italy, Vintage:*
6 *1996*}, as well as any other desired descriptive information. If a document is associated
7 with a given term, then the document is also associated with all of the terms that the given
8 term refines.

9 The classification process may proceed according to a variety of workflows.
10 Documents may be classified in series or in parallel, and the automatic and manual
11 classification steps may be performed one or more times and in any order. To improve
12 accuracy and throughput, human experts may be assigned as specialists to oversee the
13 classification task for particular subsets of the documents, or even particular attributes for
14 particular subsets of the documents. In addition, the classification and taxonomy
15 processes may be interleaved, especially as knowledge gained from one process allows
16 improvements in the other.

17 Figure 15 illustrates the stages in a possible flow for the classification process
18 250. The data acquisition step 252, that is, the collection of documents for the database,
19 may occur in several different ways. For example, a retailer with a product catalog over
20 which the search and navigation system will operate might provide a set of documents
21 describing its products as a pre-defined set. Alternatively, documents may be collected
22 from one source, e.g., one Web site, or from a number of sources, e.g., multiple Web

1 sites, and then aggregated. If the desired documents are Web pages, the documents may
2 be collected by appropriately crawling the Web, selecting documents, and discarding
3 documents that do not fit in the domain. In the data translation step 254, the collected
4 documents are formatted and parsed to facilitate further processing. In the automatic
5 classification step 256, the formatted and parsed documents are processed in order to
6 automatically associate documents with terms. In the manual classification step 258,
7 human reviewers may verify and amend the automatic classifications, thereby ensuring
8 quality control. Preferably, any rules or expectations violated in either the automatic
9 classification step 256 or the manual classification step 258 would be flagged and
10 presented to human reviewers as part of the manual classification step 258. If the
11 collection of documents is divided into domains, then there will typically be rules that
12 specify a certain minimal or preferred set of attributes used to classify documents from
13 each domain, as well as other domain-specific classification rules. When the classification
14 process is complete, each document will have a set of terms associated with it, which
15 locate the document in the collection of navigation states.

16 In Figure 16, table 180 shows a possible representation of a collection of classified
17 wine bottles. Preferably, each entry is associated with a document number 182, which
18 could be a universal identifier, a name 184, and the associated terms 186. The name is
19 preferably descriptive information that could allow the collection to be accessed via a
20 free-text search engine as well as via term-based navigation.

21 In another aspect of the invention, the knowledge base also includes a catalog of
22 canonical representations of documents. Each catalog entry represents a conceptually

1 distinct item that may be associated with one or more documents. The catalog allows
2 aggregation of profile information from multiple documents that relate to the item,
3 possibly from multiple sources. For example, if the same wine is sold by two vendors,
4 and if one vendor provides vintage and geographic location information and another
5 provides taste information, that information from the two vendors can be combined in the
6 catalog entry for that type of wine. The catalog may also improve the efficiency of the
7 classification process by eliminating duplicative profiling. In Figure 15, the catalog
8 creation step 260 associates classified documents with catalog entries, creating new
9 catalog entries when appropriate. For ease of reference, an item may be uniquely
10 identified in the catalog by a universal identifier.

11 The knowledge base may also define stores, where a store is a subcollection of
12 documents that are grouped to be retrievable at one time. For example, a particular online
13 wine merchant may not wish to display documents corresponding to products sold by that
14 merchant's competitors, even though the knowledge base may contain such documents.
15 In this case, the knowledge base can define a store of documents that does not include
16 wines sold by the merchant's competitors. In Figure 15, the store creation step 262 may
17 define stores based on attributes, terms, or any other properties of documents. A
18 document may be identified with more than one store. The knowledge base may also
19 contain attributes or terms that have been customized for particular stores.

20 In Figure 15, the export process step 264 exports information from the knowledge
21 base to another stage in the system that performs further processing necessary to generate
22 a navigable data structure.

1 **Navigation States**

2 The search and navigation system represents, explicitly or implicitly, a collection
3 of navigation states. A navigation state can be represented either by an expression of
4 terms, or by the subset of the collection of documents that correspond to the term
5 expression.

6 By way of example, types of navigation states include conjunctive navigation
7 states, disjunctive navigation states and negational navigation states. Conjunctive
8 navigation states are a special case of navigation states in which the term expression is
9 conjunctive—that is, the expression combines terms using only the AND operator.
10 Conjunctive navigation states are related by a partial order of refinement that is derived
11 from the partial order that relates the terms.

12 In one aspect of the present invention, a conjunctive navigation state has two
13 representations. First, a conjunctive navigation state corresponds to a subset of the
14 collection of documents. Second, a conjunctive navigation state corresponds to a
15 conjunctive expression of mutually incomparable terms. Figure 18 illustrates some
16 navigation states for the documents and terms based on the wine example discussed
17 above. For example, one navigation state 224 is {*Origin: South America*} (documents #1,
18 #4, #5); a second navigation state 224 is {*Type/Varietal: White AND Origin: United*
19 *States*} (documents #2, #9). The subset of documents corresponding to a conjunctive
20 navigation state includes the documents that are commonly associated with all of the
21 terms in the corresponding expression of mutually incomparable terms. At the same time,
22 the expression of mutually incomparable terms corresponding to a conjunctive navigation

1 state includes all of the minimal terms from the terms that are common to the subset of
2 documents, i.e., the terms that are commonly associated with every document in the
3 subset. A conjunctive navigation state is preferably unique and fully specified; for a
4 particular conjunctive expression of terms, or for a given set of documents, there is no
5 more than one corresponding conjunctive navigation state.

6 One way preferred to define the collection of conjunctive navigation states is to
7 uniquely identify each conjunctive navigation state by a canonical conjunctive expression
8 of mutually incomparable terms. A two-step mapping process that maps an arbitrary
9 conjunctive expression of terms to a canonical conjunctive expression of mutually
10 incomparable terms creates states that satisfy this property. In the first step of the
11 process, an arbitrary conjunctive expression of terms is mapped to the subset of
12 documents that are associated with all of those terms. Recalling that if a document is
13 associated with a given term, then the document is also associated with all of the terms
14 that the given term refines, in the second step of the process, this subset of documents is
15 mapped to the conjunctive expression of minimal terms among the terms that are
16 common to all of the documents in that document set. The result of this second step is a
17 conjunctive expression of mutually incomparable terms that uniquely identifies the
18 corresponding subset of documents, and, hence, is a canonical representation for a
19 conjunctive navigation state. By way of illustration, referring to the wine example in
20 Figure 17, the term expression {*Origin: France*} maps to the subset of documents
21 {documents #8, #11}, which in turn maps to the canonical term expression
22 {*Type/Varietal: Red AND Origin: France*}.

1 The conjunctive navigation states 222, 224, 226 are related by a partial order of
2 refinement relationships 220 derived from the partial order that relates terms. This partial
3 order can be expressed in terms of either the subsets of documents or the term expressions
4 that define a conjunctive navigation state. Expressed in terms of subsets of documents, a
5 navigation state A refines a navigation state B if the set of documents that corresponds to
6 state A is a subset of the set of documents that corresponds to state B. Expressed in terms
7 of term expressions, a conjunctive navigation state A refines a conjunctive navigation
8 state B if all of the terms in state B either are in state A or are refined by terms in state A.
9 Referring to Figure 17, the navigation state 226 corresponding to the term expression
10 *{Type/Varietal: Red AND Origin: Chile}* (document #4) refines the navigation state 224
11 corresponding to *{Origin: Chile}* (documents #4, #5). Since the refinement relationships
12 among navigation states give rise to a partial order, they are transitive and antisymmetric.
13 In the example, *{Type/Varietal: Red AND Origin: Chile}* (document #4) refines *{Origin:*
14 *Chile}* (documents #4, #5) and *{Origin: Chile}* (documents #4, #5) refines *{Origin:*
15 *South America}* (documents #1, #4, #5); therefore, *{Type/Varietal: Red AND Origin:*
16 *Chile}* (document #4) refines *{Origin: South America}* (documents #1, #4, #5). The root
17 navigation state 222 is defined to be the navigation state corresponding to the entire
18 collection of documents. The leaf navigation states 226 are defined to be those that
19 cannot be further refined, and often (though not necessarily) correspond to individual
20 documents. There can be arbitrarily many intermediate navigation states 224 between the
21 root 222 and the leaves 226. Given a pair of navigation states A and B where B refines
22 A, there can be multiple paths of intermediate navigation states 224 connecting A to B in

1 the partial order. For convenience of definition in reference to the implementation
2 described herein, a navigation state is considered to refine itself.

3 A user browses the collection of documents by visiting a sequence of one or more
4 navigation states typically starting at the root navigation state 222. In one embodiment of
5 the present invention, there are three basic modes of navigation among these states. The
6 first mode is refinement, or moving from the current navigation state to a navigation state
7 that refines it. The user can perform refinement either by adding a term through
8 conjunctive selection to the current navigation state or by refining a term in the current
9 navigation state; i.e., replacing a term with a refinement of that term. After the user adds
10 or refines a term, the new term expression can be mapped to a canonical term expression
11 according to the two-step mapping described above. The second mode is generalization,
12 or moving from the current navigation state to a more general navigation state that the
13 current state refines. The user can perform generalization either by removing a term from
14 the current navigation state or by generalizing a term in the current navigation state; i.e.,
15 replacing a current term with a term that the current term refines. After the user removes
16 or generalizes a term, the new term expression can be mapped to a canonical term
17 expression. The third mode is simply creating a query in the form of a desired term
18 expression, which again can be mapped to a canonical term expression to obtain a
19 navigation state.

20 In other embodiments of the present invention, there are additional modes of
21 navigation. In systems that support the corresponding types of navigation states, these
22 modes may include generalization of the navigation state through disjunctive selection, as

1 shown in Figure 19, as well as refinement of the navigation state through negational
2 selection, as shown in Figure 20. In general, terms can be combined using Boolean logic.
3 Although term expressions that are not conjunctive do not necessarily have canonical
4 forms, some implementations may be based on a system that uses a collection of
5 conjunctive navigation states. One implementation is based on logical equivalence rules
6 as described below.

7 **Implementation**

8 The knowledge base is transferred to a navigable data structure in order to
9 implement the present invention. The navigation states may be fully precomputed,
10 computed dynamically at run-time, or partially precomputed. A cache may be used to
11 avoid redundant computation of navigation states.

12 In preferred embodiments, the collection of conjunctive navigation states may be
13 represented as a graph—preferably, a directed acyclic multigraph with labeled edges. A
14 graph is a combinatorial structure consisting of nodes and edges, where each edge links a
15 pair of nodes. The two nodes linked by an edge are called its endpoints. With respect to
16 the present invention, the nodes correspond to conjunctive navigation states, and the
17 edges represent transitions that refine from one conjunctive navigation state to another.
18 Since refinement is directional, each edge is directed from the more general node to the
19 node that refines it. Because there is a partial order on the navigation states, there can be
20 no directed cycles in the graph, i.e., the graph is acyclic. Preferably, the graph is a
21 multigraph, since it allows the possibility of multiple edges connecting a given pair of
22 nodes. Each edge is labeled with a term. Each edge has the property that starting with

1 the term set of the more general end point, adding the edge term, and using the two-step
2 map to put this term set into canonical form leads to a refinement which results in the
3 navigation state that is the other endpoint. That is, each edge represents a refinement
4 transition between nodes based on the addition of a single term.

5 The following definitions are useful for understanding the structure of the graph:
6 descendant, ancestor, least common ancestor (LCA), proper ancestor, proper descendant,
7 and greatest lower bound (GLB). These definitions apply to the refinement partial order
8 among terms and among nodes. If A and B are terms and B refines A, then B is said to be
9 a descendant of A and A is said to be an ancestor of B. If, furthermore, A and B are
10 distinct terms, then B is said to be a proper descendant of A and A is said to be a proper
11 ancestor of B. The same definitions apply if A and B are both nodes.

12 If C is an ancestor of A and C is also an ancestor of B, then C is said to be a
13 common ancestor of A and B, where A, B, and C are either all terms or all nodes. The
14 minimal elements of the set of common ancestors of A and B are called the least common
15 ancestors (LCAs) of A and B. If no term has a pair of incomparable ancestors, then the
16 LCA of two terms—or of two nodes—is unique. For example, the LCA of *Origin:*
17 *Argentina* and *Origin: Chile* is *Origin: South America* in the partial order of terms 110 of
18 Figure 14B. In general, however, there may be a set of LCAs for a given pair of terms or
19 nodes.

20 In an implementation that fully precomputes the collection of nodes, computation
21 of the nodes in the graphs is preferably performed bottom-up.

1 The leaf nodes in the graph—that is, the nodes corresponding to leaf navigation
2 states—may be computed directly from the classified documents. Typically, but not
3 necessarily, a leaf node will correspond to a set containing a single document. The
4 remaining, non-leaf nodes are obtained by computing the LCA-closure of the leaf
5 nodes—that is, all of the nodes that are the LCAs of subsets of the leaf nodes.

6 The edges of the graph are determined according to a refinement function, called
7 the R function for notational convenience. The R function takes as arguments two nodes
8 A and B, where A is a proper ancestor of B, and returns the set of maximal terms such
9 that, if term C is in R (A, B), then refining node A with term C results in a node that is a
10 proper descendant of A and an ancestor (not necessarily proper) of B. For example, in
11 Figure 17, $R(\{Type/Varietal: Red\}, \{Type/Varietal: Merlot \text{ AND } Origin: Argentina$
12 $\text{ AND } Vintage: 1998\}) = \{Type/Varietal: Merlot \text{ AND } Origin: South America \text{ AND }$
13 $Vintage: 1998\}$. If B_1 is an ancestor of B_2 , then $R(A, B_1)$ is a subset of $R(A, B_2)$ —
14 assuming that A is a proper ancestor of both B_1 and B_2 . For example, $R(\{Type/Varietal:$
15 $Red\}, \{Type/Varietal: Red \text{ AND } Origin: South America\}) = \{Origin: South America\}$.

16 In the graph, the edges between nodes A and B will correspond to a subset of the
17 terms in R (A, B). Also, no two edges from a single ancestor node A use the same term
18 for refinement. If node A has a collection of descendant nodes $\{B_1, B_2, \dots\}$ such that term
19 C is in all of the $R(A, B_i)$, then the only edge from node A with term C goes to LCA ($B_1,$
20 B_2, \dots), which is guaranteed to be the unique maximal node among the B_i . In Figure 17,
21 for example, the edge from node $\{Type/Varietal: Red\}$ with term *Origin: South America*
22 goes to node $\{Type/Varietal: Red \text{ AND } Origin: South America\}$ rather than to that node's

1 proper descendants {*Type/Varietal: Merlot AND Origin: South America AND Vintage:*
2 *1998*} and {*Type/Varietal: Red AND Origin: Chile*}. The LCA-closure property of the
3 graph ensures the existence of a unique maximal node among the B_i . Thus, each edge
4 maps a node-term pair uniquely to a proper descendant of that node.

5 The LCA-closure of the graph results in the useful property that, for a given term
6 set S , the set of nodes whose term sets refine S has a unique maximal node. This node is
7 called the greatest lower bound (GLB) of S .

8 The graph may be computed explicitly and stored in a combinatorial data
9 structure; it may be represented implicitly in a structure that does not necessarily contain
10 explicit representations of the nodes and edges; or it may be represented using a method
11 that combines these strategies. Because the search and navigation system will typically
12 operate on a large collection of documents, it is preferred that the graph be represented by
13 a method that is scalable.

14 The graph could be obtained by computing the LCAs of every possible subset of
15 leaf nodes. Such an approach, however, grows exponentially in the number of leaf nodes,
16 and is inherently not scalable. An alternative strategy for obtaining the LCA closure is to
17 repeatedly consider all pairs of nodes in the graph, check if each pair's LCA is in the
18 graph, and add that LCA to the graph as needed. This strategy, though a significant
19 improvement on the previous one, is still relatively not scalable.

20 A more efficient way to precompute the nodes is to process the document set
21 sequentially, compute the node for each document, and add that node to the graph along
22 with any other nodes necessary to maintain LCA-closure. The system stores the nodes

1 and edges as a directed acyclic multigraph. The graph is initialized to contain a single
2 node corresponding to the empty term set, the root node. Referring to Figure 18, in
3 process 230 for inserting a new node into the graph, in step 232, for each new document
4 to be inserted into the graph that does not correspond to an existing node, the system
5 creates a new node. In step 234, before inserting the new node into the graph, the system
6 recursively generates and inserts any missing LCA nodes between the root node (or
7 ancestor node) and the new node. To ensure LCA-closure after every node insertion, the
8 system inserts the document node last, in steps 236 and 238, after inserting all the other
9 nodes that are proper ancestors of it.

10 Inserting a new node requires the addition of the appropriate edges from ancestors
11 to the node, in step 236, and to descendants out of the new node, in step 238. The edges
12 into the node are preferably determined by identifying the ancestors that have refinement
13 terms that lead into the new node and do not already have those refinement terms used on
14 edges leading to intermediate ancestors of the new node. The edges out of the node are
15 preferably determined by computing the GLB of the new node and appropriately adding
16 edges from the new node to the GLB and to nodes to which the GLB has edges.

17 The entire graph of conjunctive navigation states may be precomputed by
18 following the above procedures for each document in the collection. Computation of
19 other types of navigation states is discussed below. Precomputing of the graph may be
20 preferred where the size of the graph is manageable, or if users are likely to visit every
21 navigation state with equal probability. In practice, however, users typically visit some
22 navigation states more frequently than others. Indeed, as the graph gets larger, some

1 navigation states may never be visited at all. Unfortunately, reliable predictions of the
2 frequency with which navigation states will be visited are difficult. In addition, it is
3 generally not practical to precompute the collection of navigation states that are not
4 conjunctive, as this collection is usually much larger than the collection of conjunctive
5 navigation states.

6 An alternative strategy to precomputing the navigation states is to create indexes
7 that allow the navigation states to be computed dynamically. Specifically, each document
8 can be indexed by all of the terms that are associated with that document or that have
9 refinements associated with that document. The resulting index is generally much
10 smaller in size than a data structure that stores the graph of navigation states. This
11 dynamic approach may save space and precomputation time, but it may do so at the cost
12 of higher response times or greater computational requirements for operation. A dynamic
13 implementation may use a one-argument version of the R function that returns all
14 refinement terms from a given navigation state, as well a procedure for computing the
15 GLB of a term set.

16 It is also possible to precompute a subset of the navigation states. It is preferable
17 to precompute the states that will cost the most to compute dynamically. For example, if
18 a state corresponds to a large subset of the documents, it may be preferable to compute it
19 in advance. In one possible partial precomputation approach, all navigation states,
20 particularly conjunctive ones, corresponding to a subset of documents above a threshold
21 size may be precomputed. Precomputing a state is also preferable if the state will be
22 visited frequently. In some instances it may be possible to predict the frequency with

1 which a navigation state will be visited. Even if the frequency with which a navigation
2 state will be visited cannot be predicted in advance, the need to continually recompute
3 can be reduced by caching the results of dynamic computation. Most recently or most
4 frequently visited states may be cached.

5 As described above with respect to the interface, the system supports at least three
6 kinds of navigational operations—namely refinement, generalization, and query by
7 specifying an expression of terms. These operations may be further described in terms of
8 the graph. For query refinement, the system enumerates the terms that are on edges from
9 the node corresponding to the current navigation state. When the user selects a term for
10 refinement, the system responds by presenting the node to which that edge leads.
11 Similarly, for query generalization options, the system enumerates and selects edges that
12 lead to (rather than from) the node corresponding to the current navigation state.
13 Alternatively, query generalization may be implemented as a special case of query by
14 specifying a set of terms. For query by specifying a set of keywords, the system creates a
15 virtual node corresponding to the given term set and determines the GLB of the virtual
16 node in the graph. If no GLB is found, then there are no documents that satisfy the query.
17 Otherwise, the GLB node will be the most general node in the graph that corresponds to a
18 navigation state where all documents satisfy the query.

19 The above discussion focuses on how the system represents and computes
20 conjunctive navigation states. In some embodiments of the present invention, the user
21 interface only allows users to navigate among the collection of conjunctive navigation
22 states. In other embodiments, however, users can navigate to navigation states that are

1 not conjunctive. In particular, when the system supports navigation states that are not
2 conjunctive, the user interface may allow users to select terms disjunctively or
3 negationally.

4 If the system includes navigation states that are both conjunctive and disjunctive
5 (e.g., $\{(Products: DVDs \text{ OR } Products: Videos) \text{ AND } Director: Spike Lee\}$), then in some
6 embodiments, the system only precomputes a subset of the states, particularly if the total
7 number of navigation states is likely to be too large to maintain in memory or even
8 secondary (e.g., disk) storage. By using rules for equivalence of Boolean expressions, it
9 is possible to express any navigation state that mixes conjunction and disjunction in terms
10 of a union of conjunctive navigation states. The above example can be rewritten as
11 $\{(Products: DVDs \text{ AND } Director: Spike Lee) \text{ OR } (Products: Videos \text{ AND } Director:$
12 $Spike Lee)\}$. This approach leads to an implementation combining conjunctive and
13 disjunctive navigation states based on the above discussion, regardless of whether all,
14 some, or none of the graph of conjunctive navigation states is precomputed.

15 In preferred embodiments, disjunctive selections may be made within, but not
16 between, attributes. When determining the set of disjunctive generalizations, the system
17 does not consider other terms from the attribute of the given disjunction to be in the
18 navigation state. For example, if the navigation state is $\{Type/Varietal: Red \text{ AND}$
19 $Origin: Chile\}$ and the system is allowing the disjunctive selection of other countries of
20 origin, then the GLB and R function will be applied to the set $\{Type/Varietal: Red\}$ rather
21 than to $\{Type/Varietal: Red \text{ AND } Origin: Chile\}$. Accordingly, the other terms for the

1 attribute of country of origin that are incomparable to "Chile" become generalization
2 options for the navigation state.

3 If the system includes navigation states that use negation (e.g., {*Products: DVDs*
4 *AND Genre: Comedy AND (NOT Director: Woody Allen)*}), then the negationally
5 selected terms can be applied to navigation states as a post-process filtering operation.
6 The above example can be implemented by taking the conjunctive navigation state
7 {*Products: DVDs AND Genre: Comedy*} and applying a filter to it that excludes all
8 movies associated with the term *Director: Woody Allen*. This approach leads to an
9 implementation including negational navigation states based on the above discussion,
10 regardless of whether all, some, or none of the graph of conjunctive navigation states is
11 precomputed.

12 As with disjunction, when determining the set of negational generalizations, the
13 system does not consider other terms from the attribute of the given negation to be in the
14 navigation state. For example, if the navigation state is {*Medium: Compact Disc AND*
15 *Artist: Prince*} and the system is allowing the negational selection of other artists (e.g.,
16 {*Artist: Prince AND NOT (Artist: The Revolution)*}), then the GLB and R function will
17 be applied to the set {*Medium: Compact Disc*} rather than to {*Medium: Compact Disc*
18 *AND Artist: Prince*}.

19 Another aspect of the present invention is the interpretation of free-text search
20 queries. As discussed above, in embodiments of the present invention, a free-text query
21 may be interpreted in two ways. A single-term interpretation maps the query to an
22 individual term in the knowledge base. A multi-term interpretation maps the query to a

1 conjunction of two or more terms in the knowledge base—that is, a combination of terms
2 that corresponds to a conjunctive navigation state.

3 A free-text query may be formed of one or more words. In a preferred
4 embodiment of the invention, a single-term interpretation of a free-text query maps the
5 query to a term that either contains or is associated with a word or words in that query. A
6 query may have more than one single-term interpretation. For example, a query of
7 *computer science* might have { *Department: Computer Science Department* } and { *School:*
8 *School of Computer Science* } as single-term interpretations. For another example, a
9 query of *zinfandel* might have { *Wine Type: Zinfandel* } and { *Wine Type: White*
10 *Zinfandel* } as single-term interpretations. Various query semantics can be used to parse
11 the search query and determine the set of single-term interpretations for a given free-text
12 query. Under conjunctive query semantics, a matching term must contain all of the words
13 in the query. Under disjunctive query semantics, a matching term must contain at least
14 one of the words in the query. Under partial match query semantics, a matching term
15 must contain a subset of the words in the query; the particular rules are application-
16 dependent. It is also possible to vary the above to meet particular application needs.
17 Variations include ignoring common “stop words” such as *the* and *of*, treating related
18 words or word forms (e.g., singular and plural forms of nouns, or synonyms) as
19 equivalent, automatic spelling correction, allowing delimited phrases (i.e., two or more
20 words required to occur contiguously in a phrase), and support for negation (i.e.,
21 excluding terms that contain particular words).

1 In a preferred embodiment of the invention, a multi-term interpretation of a free-
2 text query maps the query to a conjunction of terms that either contain or are associated
3 with a word or words in that query and that correspond to a conjunctive navigation state
4 in the system. A query may have more than one multi-term interpretation. For example,
5 a query of *security books* might have {*Media Type: Books AND Subject: Computer*
6 *Security*} and {*Media Type: Books AND Subject: Financial Security*} as multi-term
7 interpretations. As with single-term interpretations, various query semantics can be used
8 to parse the query and determine the set of multi-term interpretations for a given free-text
9 query. Under conjunctive query semantics, a matching conjunction of terms must contain
10 all of the words in the query. Under partial match query semantics, a matching
11 conjunction of terms must contain a subset of the words in the query; the particular rules
12 are application-dependent. It is also possible to vary the above to meet particular
13 application needs. Variations, as discussed above, include ignoring common “stop
14 words”, treating related words or word forms as equivalent, automatic spelling correction,
15 allowing delimited phrases, and support for negation. Regardless of the query semantics
16 used, multi-term interpretations are themselves conjunctions of terms, and therefore
17 preferably correspond to conjunctive navigation states.

18 In typical embodiments, one-word queries will only have single-term
19 interpretations, while multi-word queries may have single-term interpretations, multi-term
20 interpretations, or both. For example, a query of *casual shoes* might have {*Type: Casual*
21 *Shoes*} as a single-term interpretation and {*Type: Athletic Shoes AND Merchant: Casual*
22 *Living*} as a multi-term interpretation.

1 In a preferred embodiment of the invention, a multi-term interpretation is
2 minimal—that is, the removal of any term from the interpretation would result in an
3 interpretation that no longer satisfies the query. For example, the conjunction of terms
4 {*Media Type: Books AND Subject: Computer Security*} is a minimal interpretation of the
5 query *security books*; it is not, however, a minimal interpretation of the query *computer*
6 *security*, since removing the term {*Media Type: Books*} results in the single-term
7 interpretation {*Subject: Computer Security*} that satisfies the query. For another
8 example, the conjunction of terms {*Flower Type: Red Roses AND Quantity: Dozen*} is a
9 minimal interpretation of the query *dozen red roses*; in contrast, the conjunction of terms
10 {*Flower Type: Red Roses AND Quantity: Dozen AND Color: Red*} is not a minimal
11 interpretation for this query, since removing the term {*Color: Red*} results in a minimal
12 multi-term interpretation that satisfies the query. In a preferred embodiment of the
13 invention, disjunctive query semantics are not used for multi-term interpretations. Under
14 disjunctive query semantics, all minimal interpretations of a query are single-term
15 interpretations. Single-term interpretations are always minimal, since there is only one
16 term to remove.

17 In a preferred embodiment of the invention, the computation of single-term search
18 results uses an inverted index data structure that maps words to the terms containing
19 them. Conjunctive query semantics may be implemented by computing the intersection
20 of the term sets returned by looking up each query word in the inverted index, while
21 disjunctive query semantics may be implemented by computing the union of the term
22 sets.

1 In a preferred embodiment of the invention, the computation of multi-term search
2 results uses both an inverted index data structure that maps words to the terms containing
3 them and an inverted index data structure that maps terms to the materials associated with
4 them.

5 In a preferred embodiment of the invention, the computation of multi-term search
6 results that correspond to conjunctive navigation states involves a four step procedure.
7 However, alternative procedures may be used. The steps of an algorithm 600 for
8 receiving a query and returning the multi-term search results are indicated in Figure 23.

9 Once a query is received in step 610, in the first step 620, the system determines
10 the set of terms that contain at least one word in the search query. This step is equivalent
11 to computing single-term search results using disjunctive query semantics.

12 In the second step 630, the system partitions the set of terms into equivalence
13 classes. Each equivalence class corresponds to a non-empty subset of words from the
14 query. Two terms which match the same subset of words in the query will be assigned to
15 the same equivalence class.

16 In the third step 640, the system considers the word subsets corresponding to the
17 equivalence classes from the previous step, and computes which combinations of these
18 subsets correspond to minimal interpretations of the query. A combination of subsets
19 corresponds to a minimal interpretation if its union is equal to the entire query, but the
20 removal of any of its subsets causes the union to not be equal to the entire query.

21 In the fourth step 650, the system considers, for each combination of subsets that
22 corresponds to a minimal interpretation, the set of multi-term interpretations that can be

1 obtained from terms in the corresponding equivalence classes. These multi-term
2 interpretations can be computed by enumerating, for each combination of subsets, all
3 possible ways of choosing one term from each of the equivalence classes in the
4 combination. Each resulting set of terms that corresponds to a conjunctive navigation
5 state is added to the set of search results as a single-term (if it only contains one term) or
6 multi-term interpretation. Finally in step 660, the results are returned.

7 For example, a search query of *1996 sweet red* in the wines domain obtains multi-
8 term interpretations as follows.

9 In the first step, the following terms contain at least one of the words in the query:

10 *Year: 1996*
11 *Wine Types: Sweet Wines*
12 *Flavors: Sweet*
13 *Wine Types: Appellational Red*
14 *Wine Types: Red Wines*
15 *Wineries: Red Birch*
16 *Wineries: Red Hill*

17 In the second step, there are 3 equivalence classes:

18 Terms containing *1996*
19 *Year: 1996*
20 Terms containing *sweet*
21 *Wine Types: Sweet Wines*
22 *Flavors: Sweet*
23 Terms containing *red*
24 *Wine Types: Appellational Red*

1 *Wine Types: Red Wines*

2 *Wineries: Red Birch*

3 *Wineries: Red Hill*

4 In the third step, there is 1 combination of equivalence classes that is a minimal
5 interpretation—namely, the combination of all 3 equivalence classes.

6 In the fourth step, the 8 candidates for minimal interpretations are:

7 { *Year: 1996 AND Wine Types: Sweet Wines AND Wine Types: Appellational Red* }

8 { *Year: 1996 AND Wine Types: Sweet Wines AND Wine Types: Red Wines* }

9 { *Year: 1996 AND Wine Types: Sweet Wines AND Wineries: Red Birch* }

10 { *Year: 1996 AND Wine Types: Sweet Wines AND Wineries: Red Hill* }

11 { *Year: 1996 AND Flavors: Sweet AND Wine Types: Appellational Red* }

12 { *Year: 1996 AND Flavors: Sweet AND Wine Types: Red Wines* }

13 { *Year: 1996 AND Flavors: Sweet AND Wineries: Red Birch* }

14 { *Year: 1996 AND Flavors: Sweet AND Wineries: Red Hill* }

15 Of these, the following map to conjunctive navigation states in the system and are
16 thus returned as search results:

17 { *Year: 1996 AND Wine Types: Sweet Wines AND Wineries: Red Birch* }

18 { *Year: 1996 AND Wine Types: Sweet Wines AND Wineries: Red Hill* }

19 { *Year: 1996 AND Flavors: Sweet AND Wine Types: Appellational Red* }

20 { *Year: 1996 AND Flavors: Sweet AND Wine Types: Red Wines* }

21 { *Year: 1996 AND Flavors: Sweet AND Wineries: Red Birch* }

22 { *Year: 1996 AND Flavors: Sweet AND Wineries: Red Hill* }

23 The other two minimal interpretations do not having matching documents and do
24 not map to a navigation state in the system.

1 For another example, a search query of *casual shoes* obtains multi-term
2 interpretations as follows.

3 In the first step, the following terms contain at least one of the words in the query:

4 *Type: Casual Shoes*
5 *Merchant: Casual Living*
6 *Brand: Casual Workstyles*
7 *Type: Athletic Shoes*
8 *Type: Dress Shoes*
9 *Brand: Goody Two Shoes*
10 *Merchant: Simple Shoes*

11 In the second step, there are 3 equivalence classes:

12 Terms containing *casual*
13 *Merchant: Casual Living*
14 *Brand: Casual Workstyles*
15 Terms containing *shoes*
16 *Type: Athletic Shoes*
17 *Type: Dress Shoes*
18 *Brand: Goody Two Shoes*
19 *Merchant: Simple Shoes*

20 Terms containing both *casual* and *shoes*

21 *Type: Casual Shoes*

22 In the third step, there are 2 combinations of equivalence classes that are minimal
23 interpretations. The first combination consists of the first two equivalence classes. The
24 second combination consists of the third equivalence class by itself.

25 In the fourth step, the 9 candidates for minimal interpretations are:

- 1 - {Merchant: Casual Living AND Type: Athletic Shoes}
- 2 {Merchant: Casual Living AND Type: Dress Shoes}
- 3 {Merchant: Casual Living AND Brand: Goody Two Shoes}
- 4 {Merchant: Casual Living AND Merchant: Simple Shoes}
- 5 {Brand: Casual Workstyles AND Type: Athletic Shoes}
- 6 {Brand: Casual Workstyles AND Type: Dress Shoes}
- 7 {Brand: Casual Workstyles AND Brand: Goody Two Shoes}
- 8 {Brand: Casual Workstyles AND Merchant: Simple Shoes}
- 9 {Type: Casual Shoes}

10 Of these, the following map to conjunctive navigation states in the system and are
11 thus returned as search results:

- 12 {Merchant: Casual Living AND Type: Athletic Shoes}
- 13 {Type: Casual Shoes}

14 The other minimal interpretations do not have matching documents and do not
15 map to a navigation state in the system. For example, the brand Casual Workstyles does
16 not sell Athletic Shoes in the system.

17 Another aspect of the present invention is its scalability through parallel or
18 distributed computation. One way to define scalability in a search and navigation system
19 is in terms of four problem dimensions: the number of materials in the collection, the
20 number of terms associated with each material in the collection, the rate at which the
21 system processes queries (throughput), and the time necessary to process a query
22 (latency). In this definition, a system is scalable if it can be scaled along any of these
23 four dimensions at a subquadratic cost. In other words:

- 1 1. If the number of materials in the collection is denoted by the variable n_1 and the
2 other three problem dimensions are held constant, then the resource requirements
3 are subquadratic in n_1 .
- 4 2. If the number of terms associated with each material in the collection is denoted
5 by the variable n_2 and the other three problem dimensions are held constant, then
6 the resource requirements are subquadratic in n_2 .
- 7 3. If the number of queries that the system processes per second (i.e., the throughput)
8 is denoted by the variable n_3 and the other three problem dimensions are held
9 constant, then the resource requirements are subquadratic in n_3 .
- 10 4. If the time necessary to process a query (i.e., the latency) is denoted by the
11 variable n_4 and the other three problem dimensions are held constant, then the
12 resource requirements are subquadratic in $1/n_4$.
- 13 Preferably, these resource requirements would be not only subquadratic, but
14 linear. Also included within the concept of scalability, there is an allowance for overhead
15 in creating a network of distributed resources. Typically, this overhead will be
16 logarithmic, since the resources may be arranged in a hierarchical configuration of
17 bounded fan-out.
- 18 In some embodiments, the present invention surmounts the limitations of a single
19 computational server's limited resources by allowing for distributing the task of

1 computing the information associated with a navigation state onto a hierarchy of multiple
2 computational servers that act in parallel.

3 One insight that drives this aspect of the present invention is that it is possible to
4 partition the collection of materials among multiple "slave" servers, all of which
5 implement the single-server algorithm for multidimensional navigation, and then to have
6 a "master" server compute navigation states by passing requests onto the set of slave
7 machines and combining the responses. From the outside, the collection of servers
8 appears to act like a single server, but with far greater computational resources than
9 would be possible on a single computational device. Indeed, the distinction between
10 master and slave servers is arbitrary; a slave server can itself have slaves, thus creating a
11 nested hierarchy of servers. Such nesting is useful when the number of slaves exceeds the
12 fan-out capability of a single master server. An exemplary embodiment of such a system
13 is illustrated in Figure 24. In the hierarchical arrangement 500, a master server 520
14 works with slave servers 530, 540. In the hierarchical arrangement shown, slave servers
15 530 are in turn master servers with respects to slave servers 540. The search and
16 navigation results are made available to a user on a terminal 510 through a user interface
17 in accordance with the present invention.

18 The collection of materials may be partitioned by splitting (arbitrarily or
19 otherwise) the materials into disjoint subsets, where each subset is assigned to its own
20 server. The subsets may be roughly equal in size, or they might vary in size to reflect the
21 differing computational resources available to each server.

1 The algorithm for distributing the task of computing the information associated
2 with a navigation state includes three steps. The steps of the algorithm are indicated in
3 Figure 24. In the first step, the query, which is a request for a valid navigation state, is
4 submitted to the master server 520, which forwards the query to each of the slave servers
5 530. If the servers are nested, the requests are forwarded through the hierarchy of servers
6 500 until they reach the leaf servers 540 in the hierarchy. In the second step, each slave
7 server 530, 540 processes the query independently, based on the subset of the collection
8 of materials that is in its partition. In the third step, the master server 520 combines the
9 responses from the slave servers to produce a response for the original query. The master
10 server 520 returns the response to the terminal 510.

11 The master server receives the original request and farms it out to the slave
12 servers. Thus, in preferred embodiments, the only computation performed by the master
13 server is to combine the results from the slave servers. Each slave server that receives a
14 request computes the navigation state based on the subset of the collection assigned to it.
15 The computation may involve any combination of conjunction, disjunction, and negation.

16 The master server, in contrast, only performs a combination step. The
17 combination step involves producing a valid navigation state, including documents and
18 corresponding refinement options, from the responses from the slave servers. Since the
19 collection of materials has been partitioned into disjoint subsets, the documents identified
20 by each of the slave servers can be combined efficiently as a disjoint union. Combining
21 the various refinement options returned by each of the slave servers may require
22 additional processing, as described below.

1 The slave servers all process the same query, but on different partitions of the
2 collection of materials. They will generally return different sets of refinement options
3 because a set of refinement options that is valid for one partition may be invalid for
4 another. If the different sets are disjoint, and if the refinement options involve terms that
5 do not themselves have refinement relationships, then the combination is a disjoint union.

6 Typically, there will be some overlap among the different sets of refinement
7 options returned by each slave server. If the sets are not disjoint, duplicates can be
8 eliminated in this combination step.

9 When there are refinement relationships among the terms that are refinement
10 options returned by the slave servers, the combination algorithm computes, for every set
11 of related terms, the least common ancestor or ancestors (LCA) of the terms, as defined
12 by the partial order among the terms. One algorithm for combining the refinement
13 options is outlined in Figure 25. In step 552, the master server receives and takes the
14 union of all of the terms, x_1, x_2, \dots, x_n , returned as refinement options for the navigation
15 state from the slave servers. In step 554, the master server computes the set of ancestors
16 A_1, A_2, \dots, A_n , for each of the terms, x_1, x_2, \dots, x_n , respectively. In step 556, the master
17 server computes the intersection A of all of the sets of ancestors, A_1, A_2, \dots, A_n . In step
18 558, the master server computes the set M of minimal terms in A . The set M , formed of
19 the least common ancestors of the terms x_1, x_2, \dots, x_n , returned by the slave servers, is the
20 set of refinement options corresponding to the result navigation state. This combination
21 procedure is applied whether the refinement options are conjunctive, disjunctive, or
22 negational.

1 In summary, the master server receives a request for a navigation state, forwards
2 this request to each of the slave servers, combines their results with a union operation,
3 and then computes, for every set of terms, the least common ancestor or ancestors of the
4 set.

5 There are at least two ways to compute the LCA of the terms. One approach is to
6 store all non-leaf terms on the master server. This strategy is reasonably memory
7 efficient, since, in practice, most of the terms are leaves (minimal elements) in the partial
8 order. A second approach is to include the ancestors when returning the terms that are
9 refinements. This approach saves memory at the expense of increasing the size of the
10 data being transferred. The latter overhead is reasonable, since, in practice, a term
11 typically has very few ancestors.

12 The task of computing results for a free-text search query may also be distributed.
13 In the arrangement described above, for example, the master can simply compute the
14 union of the free-text search results returned by the slave servers. This approach applies
15 to both single-term and multi-term search under both conjunctive and disjunctive query
16 semantics. More complex approaches may be used to accommodate customized query
17 semantics.

18 The search and navigation system of the present invention allows information
19 providers to overlay a search and navigation system over any collection of documents.
20 The knowledge base aspect and the search and navigation aspect of the invention can be
21 performed independently by different providers, and information providers may outsource
22 these functions to separate entities. Similarly, a generated knowledge base may be

1 imported by a search and navigation specialist. Information providers may also outsource
2 this search and navigation requirement to a search and navigation system provider. A
3 search and navigation system provider could charge customers a license fee for the
4 system independent of the amount of its usage. Alternatively, a search and navigation
5 system provider could charge customers on a per-click basis, a per-purchase basis if
6 products are available via the system, or per-transaction generated from a click through
7 the search and navigation system. A search and navigation system provider could also
8 function as an aggregator -- compiling records from a number of sources, combining them
9 into a global data set, and generating a search and navigation system to search and
10 navigate the data set. The search and navigation system can be implemented as software
11 provided on a disk, on a CD, in memory, etc., or provided electronically (such as over the
12 Internet).

13 A search and navigation system in accordance with the present invention may also
14 enhance user profiling capability and merchandising capability. The search and
15 navigation system may maintain a profile of users based on the users' selections,
16 including the particular paths selected to explore the collection of navigation states.
17 Using the knowledge base, the system may also infer additional information regarding the
18 users' preferences and interests by supplementing the selection information with
19 information regarding related documents, attributes and terms in the knowledge base.
20 That information may be used to market goods and services related to the documents of
21 interest to the user.

1 The foregoing description has been directed to specific embodiments of the
2 invention. The invention may be embodied in other specific forms without departing
3 from the spirit and scope of the invention. The embodiments, figures, terms and
4 examples used herein are intended by way of reference and illustration only and not by
5 way of limitation. The scope of the invention is indicated by the appended claims and all
6 changes that come within the meaning and scope of equivalency of the claims are
7 intended to be embraced therein.

8 We claim:

9

- 1 1. A search and navigation system for a set of materials, comprising:
2 a plurality of attributes characterizing the materials;
3 a plurality of values describing the materials, wherein each of the values
4 has an association with at least one of the attributes and each association defines an
5 attribute-value pair;
6 a plurality of navigation states, wherein each navigation state corresponds
7 to a particular expression of attribute-value pairs and to a particular subset of the
8 materials; and
9 a search interface, the search interface including a free-text search tool for
10 accepting free-text queries, the search interface being adapted to generate multi-term
11 interpretations of free-text queries, a multi-term interpretation including a conjunction of
12 attribute-value pairs that corresponds to a navigation state, the search interface providing a
13 display of a set of search results for a query, the set of search results including multi-term
14 interpretations.
- 15 2. The search and navigation system of claim 1, wherein the multi-term
16 interpretations of the free-text query are minimal.
- 17 3. The search and navigation system of claim 1, wherein the search interface
18 supports conjunctive query semantics.
- 19 4. The search and navigation system of claim 1, wherein the search interface
20 supports disjunctive query semantics.
- 21 5. The search and navigation system of claim 1, wherein the search interface
22 supports customized query semantics.

- 1 6. The search and navigation system of claim 1, wherein the search interface
2 ignores stop words in the free-text query.
- 3 7. The search and navigation system of claim 1, wherein the search interface
4 treats syntactically related words as equivalent.
- 5 8. The search and navigation system of claim 1, wherein the search interface
6 treats semantically related words as equivalent.
- 7 9. The search and navigation system of claim 1, wherein the search interface
8 performs automatic spelling corrections.
- 9 10. The search and navigation system of claim 1, wherein the search interface
10 supports the specification of delimited phrases.
- 11 11. The search and navigation system of claim 1, wherein the search interface
12 supports constraining the set of search results to the subset of materials in the current
13 navigation state where the free-text query is accepted.
- 14 12. The search and navigation system of claim 1, further including a profile
15 for each of the materials in the set of materials, the profile including descriptive
16 information, the free-text search tool enabling searching the descriptive information in the
17 profiles.
- 18 13. The search and navigation system of claim 1, the search interface further
19 including a full-text search tool for searching the set of materials.
- 20 14. The search and navigation system of claim 1, wherein the set of search
21 results is organized by attribute.

1 15. - The search and navigation system of claim 1, wherein the set of search
2 results further includes navigation options to the navigation states corresponding to the set
3 of search results. .

4 16. The search and navigation system of claim 1, further including a first
5 inverted index relating words to attribute-value pairs and a second inverted index relating
6 attribute-value pairs to materials.

7 17. The search and navigation system of claim 1, further comprising a
8 navigation interface, the navigation interface including a guided navigation tool
9 providing a set of navigation options from the current navigation state to other navigation
10 states, each navigation option providing a direct path to one of the other navigation states.

11 18. A search and navigation system for a set of materials, comprising:
12 a plurality of attributes characterizing the materials;
13 a plurality of values describing the materials, wherein each of the values has an
14 association with at least one of the attributes and each association defines an attribute-
15 value pair;

16 a plurality of navigation states, wherein each navigation state corresponds to a
17 particular expression of attribute-value pairs and to a particular subset of the materials;
18 and

19 a search interface, the search interface including a free-text search tool for
20 accepting free-text queries, the search interface being adapted to generate single-term and
21 multi-term interpretations of free-text queries, a single-term interpretation including an
22 attribute-value pair that corresponds to a navigation state, and a multi-term interpretation

1 including a conjunction of attribute-value pairs that corresponds to a navigation state, the
2 search interface providing a display of a set of search results for a query, the set of search
3 results including single-term interpretations or multi-term interpretations or both.

4 19. The search and navigation system of claim 1, wherein the multi-term
5 interpretations of the free-text query are minimal.

6 20. The search and navigation system of claim 18, wherein the search interface
7 supports conjunctive query semantics.

8 21. The search and navigation system of claim 18, wherein the search interface
9 supports disjunctive query semantics.

10 22. The search and navigation system of claim 18, wherein the search interface
11 supports customized query semantics.

12 23. The search and navigation system of claim 18, wherein the search interface
13 ignores stop words in the free-text query.

14 24. The search and navigation system of claim 18, wherein the search interface
15 treats syntactically related words as equivalent.

16 25. The search and navigation system of claim 18, wherein the search interface
17 treats semantically related words as equivalent.

18 26. The search and navigation system of claim 18, wherein the search interface
19 performs automatic spelling corrections.

20 27. The search and navigation system of claim 18, wherein the search interface
21 supports the specification of delimited phrases.

1 28. The search and navigation system of claim 18, wherein the search interface
2 supports constraining the set of search results to the subset of materials in the current
3 navigation state where the free-text query is accepted.

4 29. The search and navigation system of claim 18, wherein the set of search
5 results is organized by attribute.

6 30. The search and navigation system of claim 18, wherein the set of search
7 results further includes navigation options to the navigation states corresponding to the set
8 of search results.

9 31. The search and navigation system of claim 18, further including a first
10 inverted index relating words to attribute-value pairs and a second inverted index relating
11 attribute-value pairs to materials.

12 32. The search and navigation system of claim 18, further comprising a
13 navigation interface, the navigation interface including a guided navigation tool
14 providing a set of navigation options from the current navigation state to other navigation
15 states, each navigation option providing a direct path to one of the other navigation states.

16 33. A search and navigation system for a set of materials, comprising:
17 a plurality of attributes characterizing the materials;
18 a plurality of values describing the materials, wherein each of the values has an
19 association with at least one of the attributes and each association defines an attribute-
20 value pair, and wherein some of the attribute-value pairs refine other of the attribute-value
21 pairs;

1 a plurality of navigation states, wherein each navigation state corresponds to a
2 particular expression of attribute-value pairs and to a particular subset of the materials;
3 a navigation interface, the interface providing a plurality of transitions, each
4 transition providing a direct path between two of the navigation states, wherein each
5 transition represents a change from the expression of attribute-value pairs corresponding
6 to an originating navigation state to the expression of attribute-value pairs corresponding
7 to a destination navigation state, wherein a series of one or more transitions provides a
8 path between any two navigation states, there being more than one path between at least a
9 first of the navigation states and a second of the navigation states; and
10 a search interface, the interface including a free-text search tool for accepting free-
11 text queries, the interface being adapted to generate multi-term interpretations for free-
12 text queries, a multi-term interpretation including a conjunction of attribute-value pairs
13 that corresponds to a navigation state, the interface providing a set of search results
14 including multi-term interpretations for a free-text query.

15 34. A method for enabling a user to search a set of materials, a plurality of
16 attributes characterizing the materials, a plurality of values describing the materials, each
17 of the values having an association with at least one of the attributes, each association
18 defining an attribute-value pair, comprising the steps of:
19 defining a plurality of navigation states, each navigation state corresponding to a
20 particular expression of attribute-value pairs and to a particular subset of the materials;
21 receiving a free-text query;

1 generating a result set for the free-text query, including computing multi-term
2 interpretations of the free-text query; and
3 providing a display of the result set.

4 35. The method of claim 34, wherein the multi-term interpretations are
5 minimal.

6 36. The method of claim 34, the step of generating the result set further
7 including computing single-term interpretations of the free-text query.

8 37. The method of claim 34, wherein the step of generating a result set uses
9 conjunctive query semantics.

10 38. The method of claim 34, wherein the step of generating a result set uses
11 disjunctive query semantics.

12 39. The method of claim 34, wherein the step of generating a result set uses
13 partial match query semantics.

14 40. The method of claim 34, wherein the step of generating a result set treats
15 syntactically related words as equivalent.

16 41. The method of claim 34, wherein the step of generating a result set treats
17 semantically related words as equivalent.

18 42. A method determining results for a query including a plurality of words
19 directed to a set of materials, , a plurality of attributes characterizing the materials, a
20 plurality of values describing the materials, each of the values having an association with
21 at least one of the attributes, each association defining an attribute value pair, the
22 materials and the attribute-value pairs defining navigation states, each navigation state

1 corresponding to a particular expression of attribute-value pairs and to a particular subset
2 of the materials, comprising the steps of:

3 computing the set of corresponding attribute value-pairs containing at least one of
4 the plurality of words;

5 computing the set of equivalence classes of the set of corresponding attribute-
6 value-pairs;

7 computing the set of minimal conjunctions of the equivalence classes; and

8 computing for each conjunction of the equivalence classes in the set of minimal
9 conjunctions the set of corresponding single-term or multi-term interpretations that
10 contain exactly one attribute-value pair from each equivalence class in the conjunction of
11 equivalence classes and that correspond to non-empty navigation states.

12 43. A computer program product, residing on a computer readable medium,
13 for use in searching a set of materials, in which the materials are characterized by a
14 plurality of attributes, and the materials are described by a plurality of values, each of the
15 values having an association with at least one of the attributes, each association defining
16 an attribute-value pair, and in which a plurality of navigation states are defined, each
17 navigation state corresponding to a particular expression of attribute-value pairs and to a
18 particular subset of the materials, the computer program product comprising instructions
19 for causing a computer to:

20 receive a free-text query;

21 generate single-term and multi-term interpretations of the query, a single term

22 interpretation including an attribute-value pair that corresponds to a navigation state, a

- 1 multi-term interpretation including a conjunction of attribute-value pairs that corresponds
- 2 to a navigation state;
- 3 return a set of search results for the query, the set of search results including
- 4 single-term interpretations or multi-term interpretations or both.
- 5

opti-wine.com 1/25

Keyword Search Find

Begin Your Search... 24

22-Wine Types ▶ 24

22-Appellations ▶ 24

22-Wineries ▶ 24

22-Year ▶ 24

22-Special Designations ▶ 24

22-Flavors ▶ 24

22-Price Range ▶ 24

22-Regions ▶ 24

22-Wine Spectator Rating ▶ 24

22-Body & Characteristics ▶ 24

Your Selection Contains 21044 Items... 44

Page: 1 2 3 4 5 6 7 8 9 10 >>

Displaying results 1-10 30

Next > 41

price: \$13.00
score: 90-94 42

Chardonnay Monterey County
Bold, rich and spicy, with layers of complex pear, toast, honey and vanilla flavors that are intense and concentrated, with a long, full finish. Delicious now. (12000 cases produced)

price: \$15.00
score: 90-94 42

Chardonnay Monterey County
A bold, ripe and full-bodied white from California that offers lots of rich pear, spice, honey flavors, all presented with a light shading of hazelnut. This has a sense of elegance and grace that goes on through the finish. (22000 cases produced)

price: \$30.00
score: 90-94 42

Marinus Carmel Valley
Young, tight and well focused, with rich, complex flavors of spicy currant, cedar, leather, anise and berry at the core, it unfolds slowly to reveal some exotic spice and mineral notes. Given the level of intensity, it's best to cellar this one unt

price: \$10.00
score: 90-94 42

Sauvignon Blanc Monterey County
Bright and pure, pouring out its generous pear, pineapple and citrus flavors. An incredible value in a California white that's fresh and lively through the long finish. Delicious now. (2700 cases produced)

price: \$17.00
score: 90-94 42

Chardonnay Monterey County
A big, ripe Chardonnay, with an abundance of rich pear, citrus, oak and spice notes. Turns smooth and spicy on the finish, where the flavors fan out. (14676 cases produced)

price: \$12.00
score: 90-94 42

Sauvignon Blanc Monterey County
Smooth, rich and buttery, a spicy wine with generous layers of pear, honey and exotic tropical fruit character sneaking in on the silky finish. Ready now. (2100 cases produced)

price: \$18.00
score: 90-94 42

Chardonnay Monterey County
Distinct for its bright citrus, especially lemony, flavors, this well-crafted white also offers touches of pear, spice, earth and oak, holding its focus while gaining nuances of oak and hazelnut. Delicious. Drink now through 2001. (35500 cases produced)

price: \$9.00
score: 80-89 42

Merlot Monterey
Ripe plum and black cherry here, with touches of charry oak and spice on the finish. Drink now. (4500 cases produced)

price: \$14.00
score: 90-94 42

Chardonnay Monterey County La Reina Vineyard
Rich in texture and full of fruit and butter flavors. The oak is evident, but there are ample pear, apricot, butterscotch and spice for complexity. Well-rounded in the mouth and well-balanced with acidity, making the flavors vivid and the feel

20

10

FIG. 1

opti-wine.com

2/25

Begin Your Search...

- 22 - Wine Types ▾
- 22 - Appellations ▾
- Wineries ▾
- Year ▾
- Special Designations ▾
- Flavors ▾
- Price Range ▾
- 22 - Regions ▾
- Wine Spectator Rating ▾
- Body & Characteristics ▾

Your Selection Contains 21044 Items...

Page: 1 2 3 4 5 6 7 8 9 10 >>

24

Chardonnay Monterey County

Bold, rich and spicy, with layers of complex pear, toast, honey and vanilla flavors that are intense and concentrated, with a long, full finish. Delicious now. (12000 cases produced)

price: \$13.00
score: 90-94

Chardonnay Monterey County

Bodied white from lots of rich pear, all presented with a nut. This has a sense that goes on 2000 cases

price: \$15.00
score: 90-94

Chardonnay Monterey County

ey focused, with rich, spicy currant, cedar, leather, anise and berry at the core, it unfolds slowly to reveal some exotic spice and mineral notes. Given the level of intensity, it's best to cellar this one unt

price: \$30.00
score: 90-94

Sauvignon Blanc Monterey County

Bright and pure, pouring out its generous pear, pineapple and citrus flavors. An incredible value in a California white that's fresh and lively through the long finish. Delicious now. (2700 cases produced)

price: \$10.00
score: 90-94

Chardonnay Monterey County

A big, ripe Chardonnay, with an abundance of rich pear, citrus, oak and spice notes. Turns smooth and spicy on the finish, where the flavors fan out. (14676 cases produced)

price: \$17.00
score: 90-94

Sauvignon Blanc Monterey County

Smooth, rich and buttery, a spicy wine with generous layers of pear, honey and exotic tropical fruit character sneaking in on the silky finish. Ready now. (2100 cases produced)

price: \$12.00
score: 90-94

Chardonnay Monterey County

Distinct for its bright citrus, especially lemony, flavors, this well-crafted white also offers touches of pear, spice, earth and oak, holding its focus while gaining nuances of oak and hazelnut. Delicious. Drink now through 2001. (35500 cases produced)

price: \$18.00
score: 90-94

Merlot Monterey

Ripe plum and black cherry here, with touches of charry oak and spice on the finish. Drink now. (4500 cases produced)

price: \$9.00
score: 80-89

Chardonnay Monterey County La Reina Vineyard

Rich in texture and full of fruit and butter flavors. The oak is evident, but there are ample pear, apricot, butterscotch and spice for complexity. Well-rounded in the mouth and well-balanced with acidity, making the flavors vivid and the feel

price: \$14.00
score: 90-94

Keyword Search

Find

Displaying results 1-10

Next >

20

10

FIG. 2

3/25

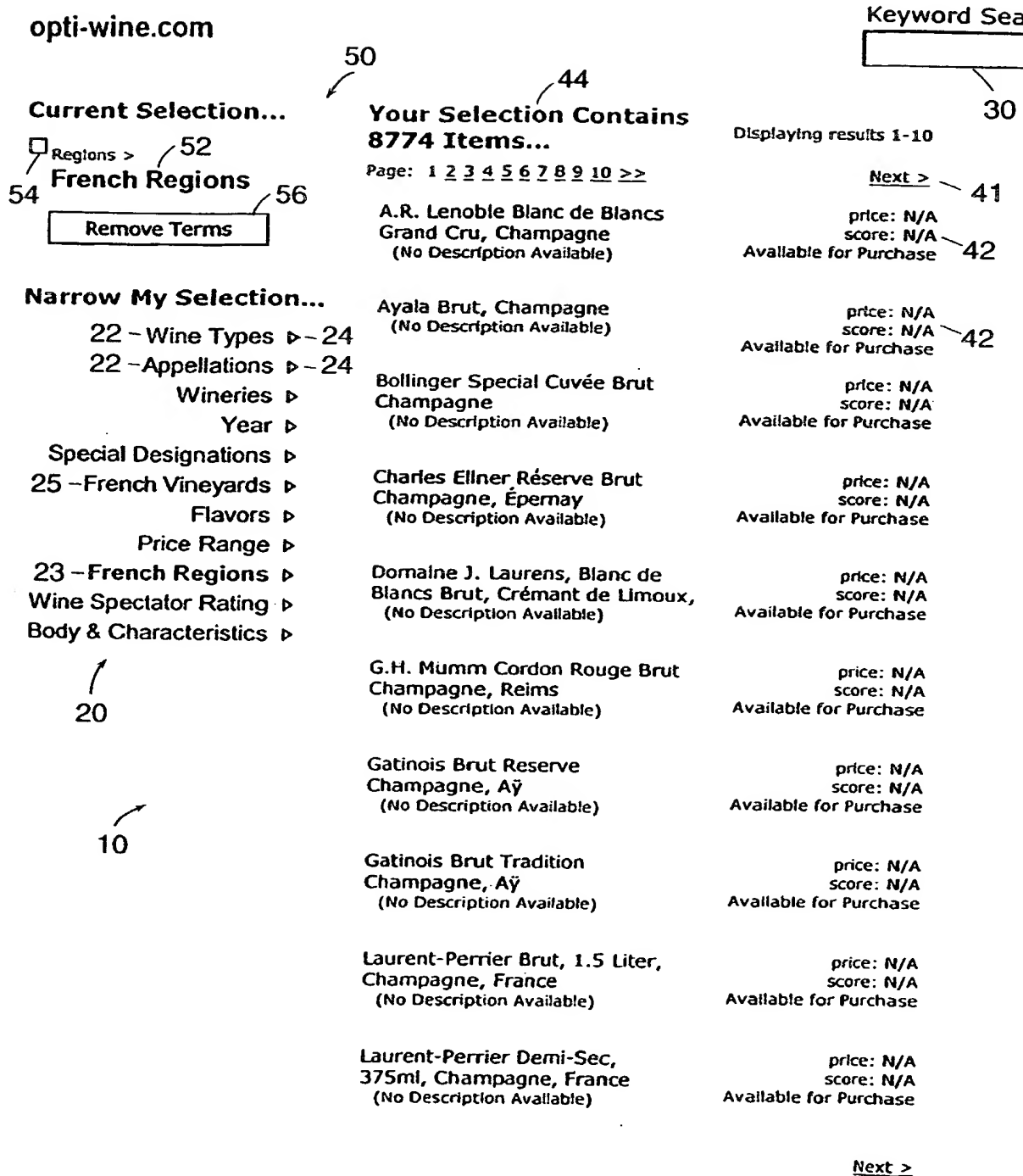


FIG. 3

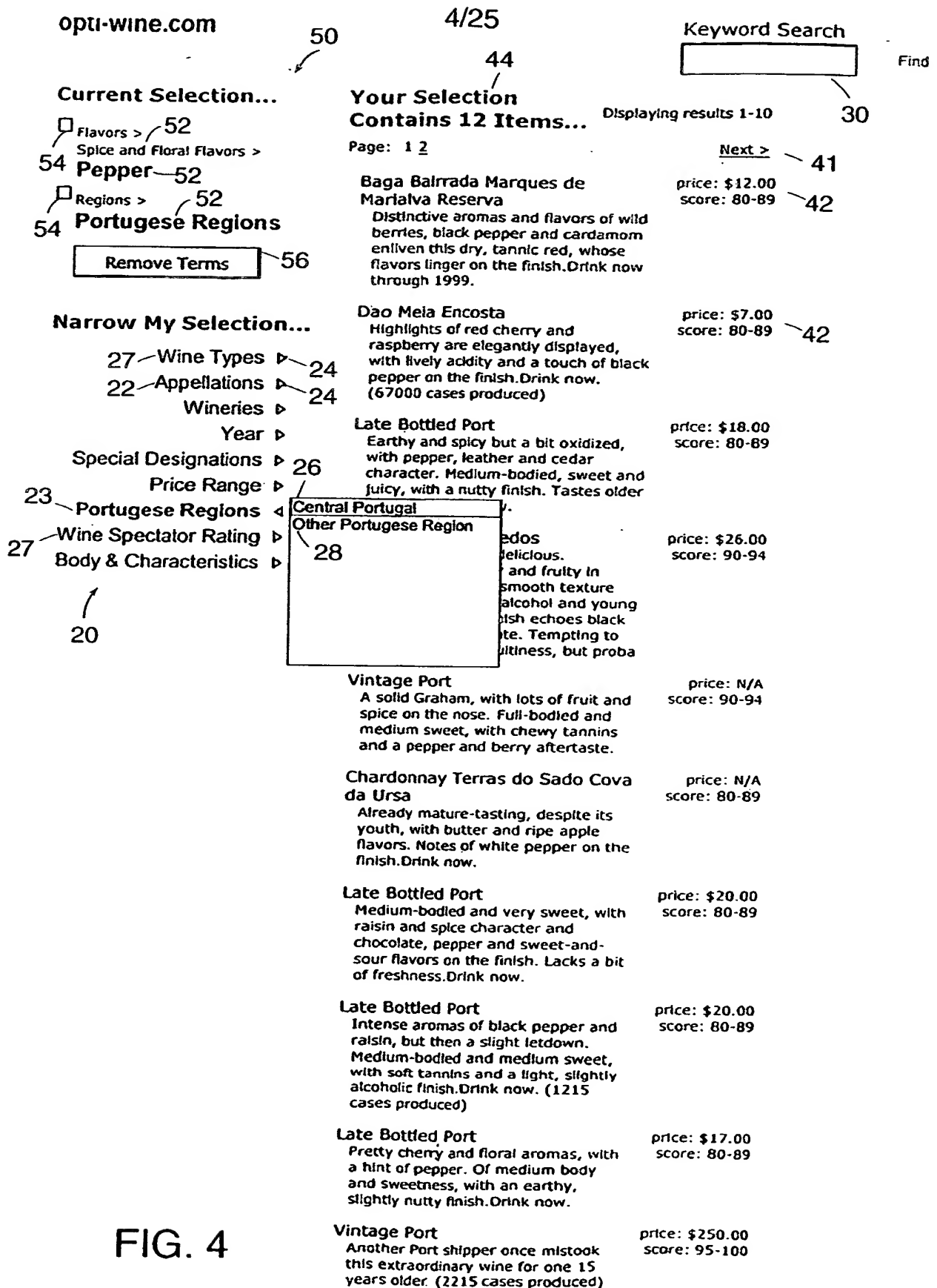


FIG. 4

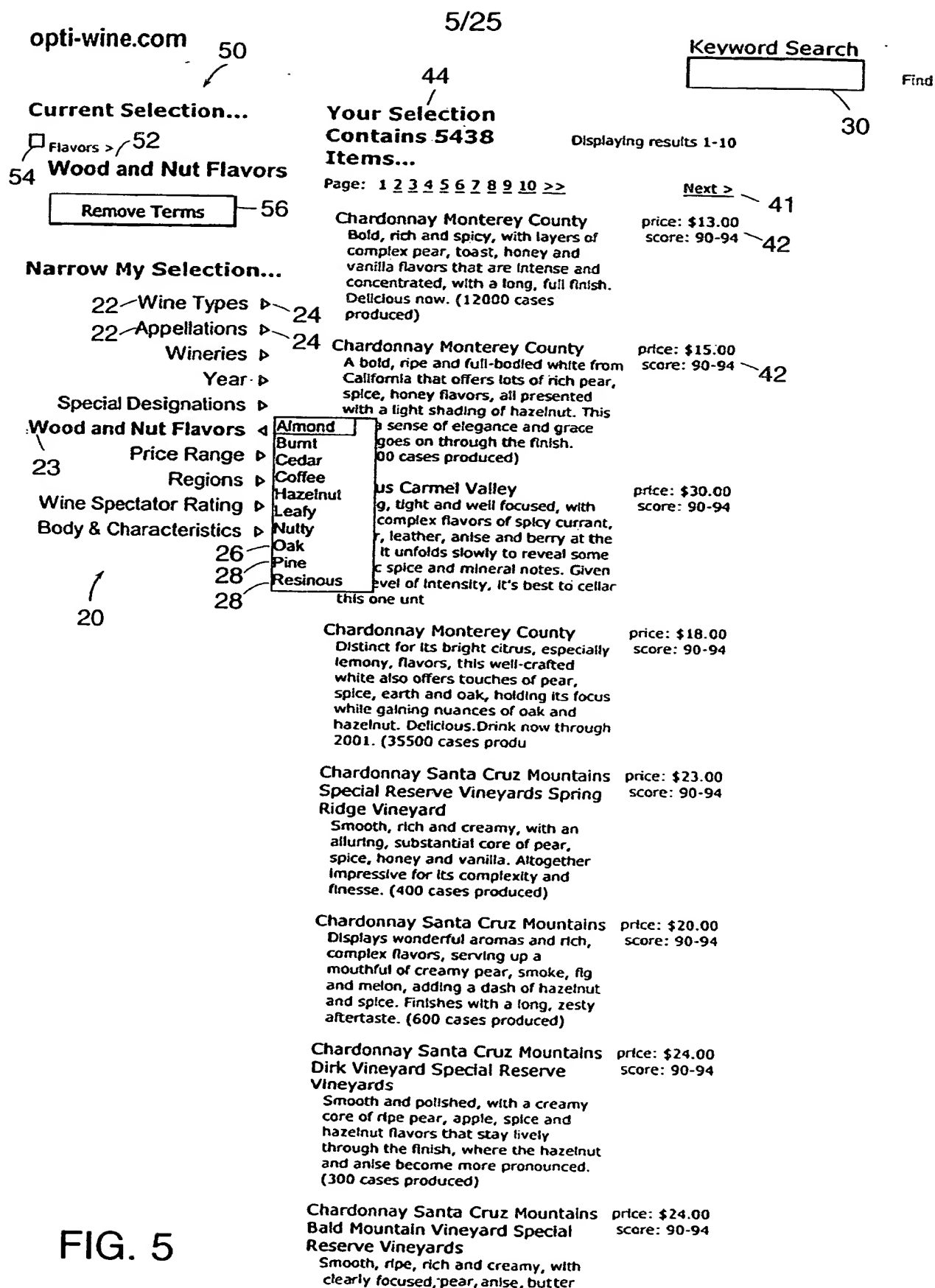


FIG. 5

6/25

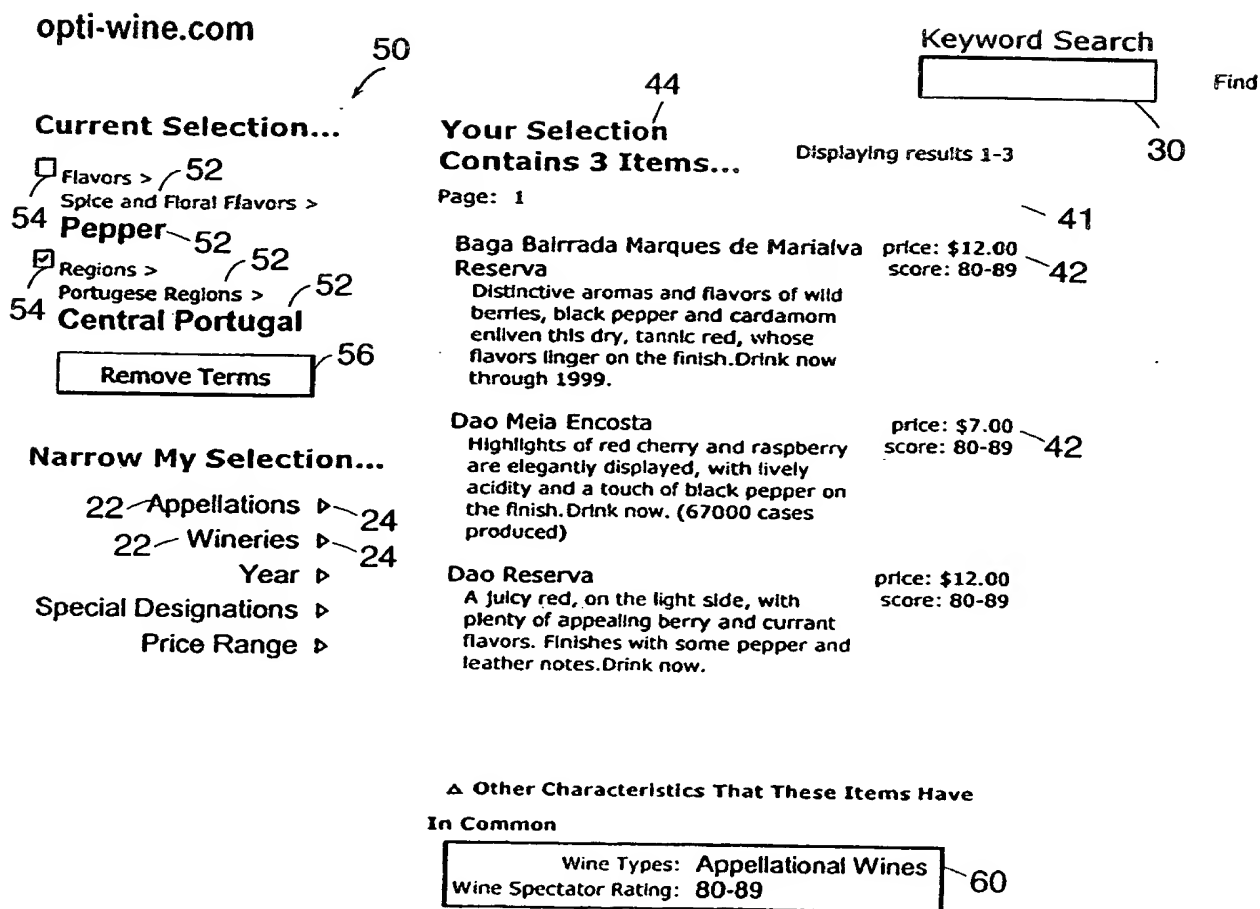


FIG. 6

7/25

opti-wine.com

Keyword Search

Find

Begin Your Search...

Your Selection

Contains 2337

Displaying results 1-10

Items...

28

22 Wine Types

22 Appellations

Wineries

Year

Special Designations

Flavors

Price Range (Wine)

Regions

Wine Spectator Range

Body & Characteristics

Adelaide Hills

Alexander Valley — 29

Alicante — 28

Alsace

Alsace Grand Cru — 28

Amarone della Valpolicella — 28

Anderson Valley — 28

Arroyo Grande Valley — 28

Arroyo Seco — 28

Arzig — 28

Next >

price: \$3.00

score: N/A

price: \$15.00

score: 90-94

price: \$105.00

score: 90-94

Cabernet Sauvignon Napa Valley

Awkward in aroma when first poured, but it has plenty of vigor in the firm tannins and deep flavors of cherry, tomato and spice. By the end of the tasting, it had blossomed into a well-aged, harmonious wine. Drink now.--Chappellet Cabernet vertical.

Zinfandel Paso Robles Dusi Ranch
(No Description Available)

price: \$9.00

score: N/A

Cabernet Sauvignon Napa Valley

price: \$95.00

score: 90-94

An outstanding wine from a great vintage for California Cabernet. A big bouquet of meaty, herbal, toasty aromas gives way to lively fruit flavors and a firm, fresh texture. Drink now through 1996.--Chappellet Cabernet vertical.

Petite Sirah Napa Valley
(No Description Available)

price: N/A

score: 90-94

Cabernet Sauvignon Napa Valley Red Rock Terrace

price: N/A

score: 90-94

Very complex, with a broad range of earthy currant, plum, berry, sage and spice flavors. Long, intricate, lingering aftertaste.--Diamond Creek vertical.

Cabernet Sauvignon Napa Valley Volcanic Hill

price: N/A

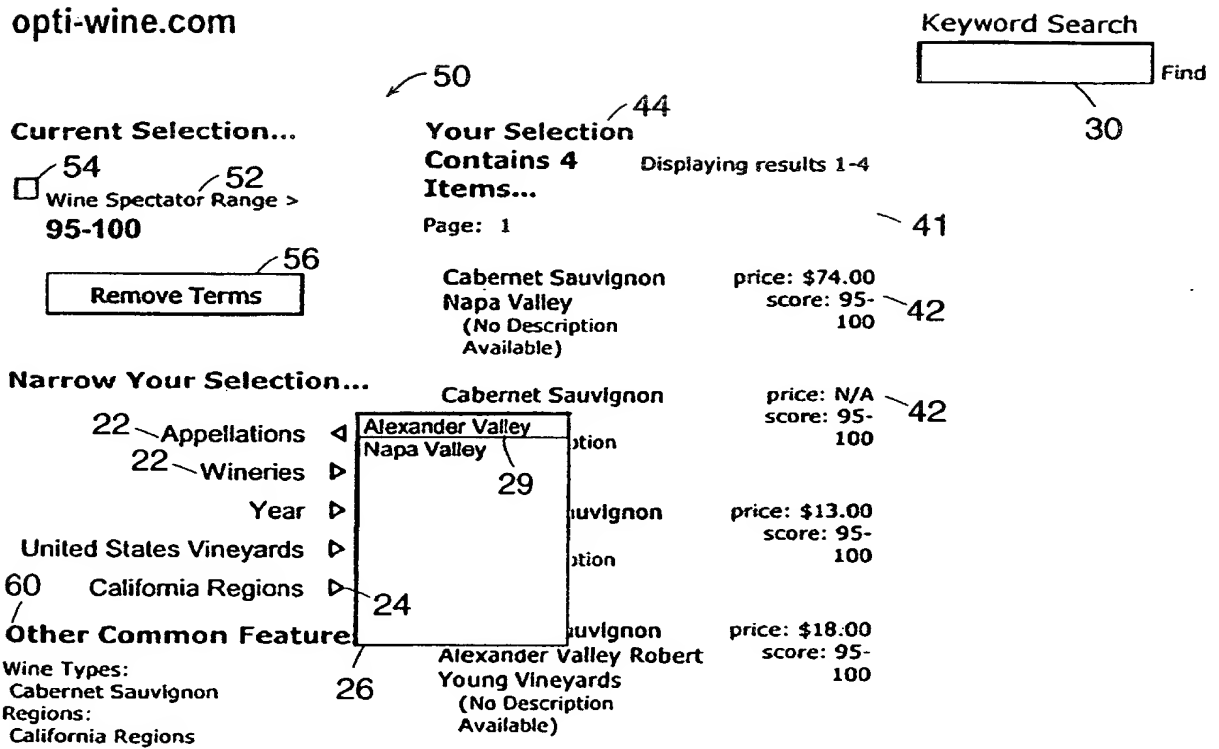
score: 80-89

Austere, with a thin band of mature Cabernet flavors. Less complex, flavorful and interesting than the 1972.--Diamond Creek vertical.

FIG. 7

8/25

opti-wine.com

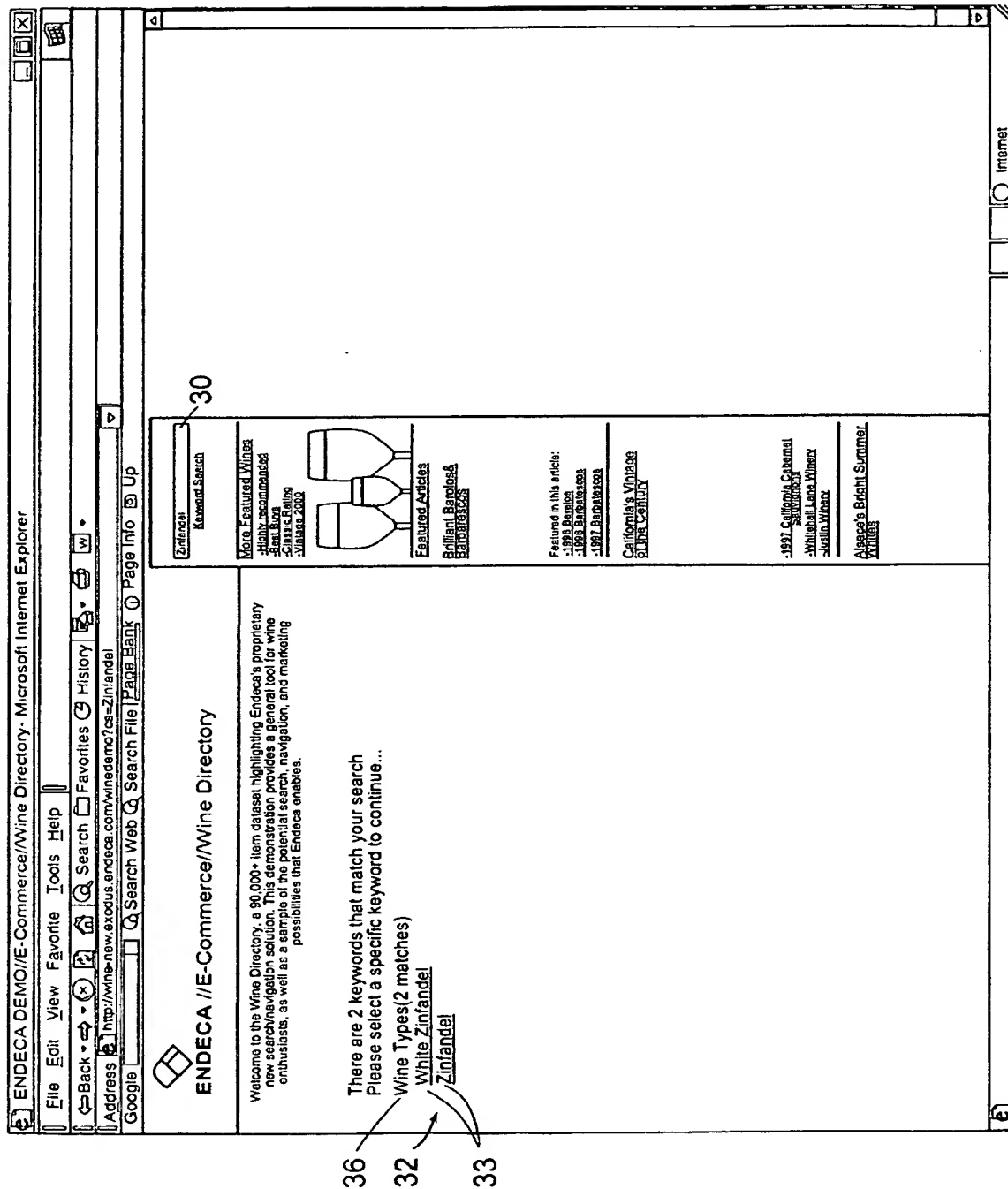


10

FIG. 8

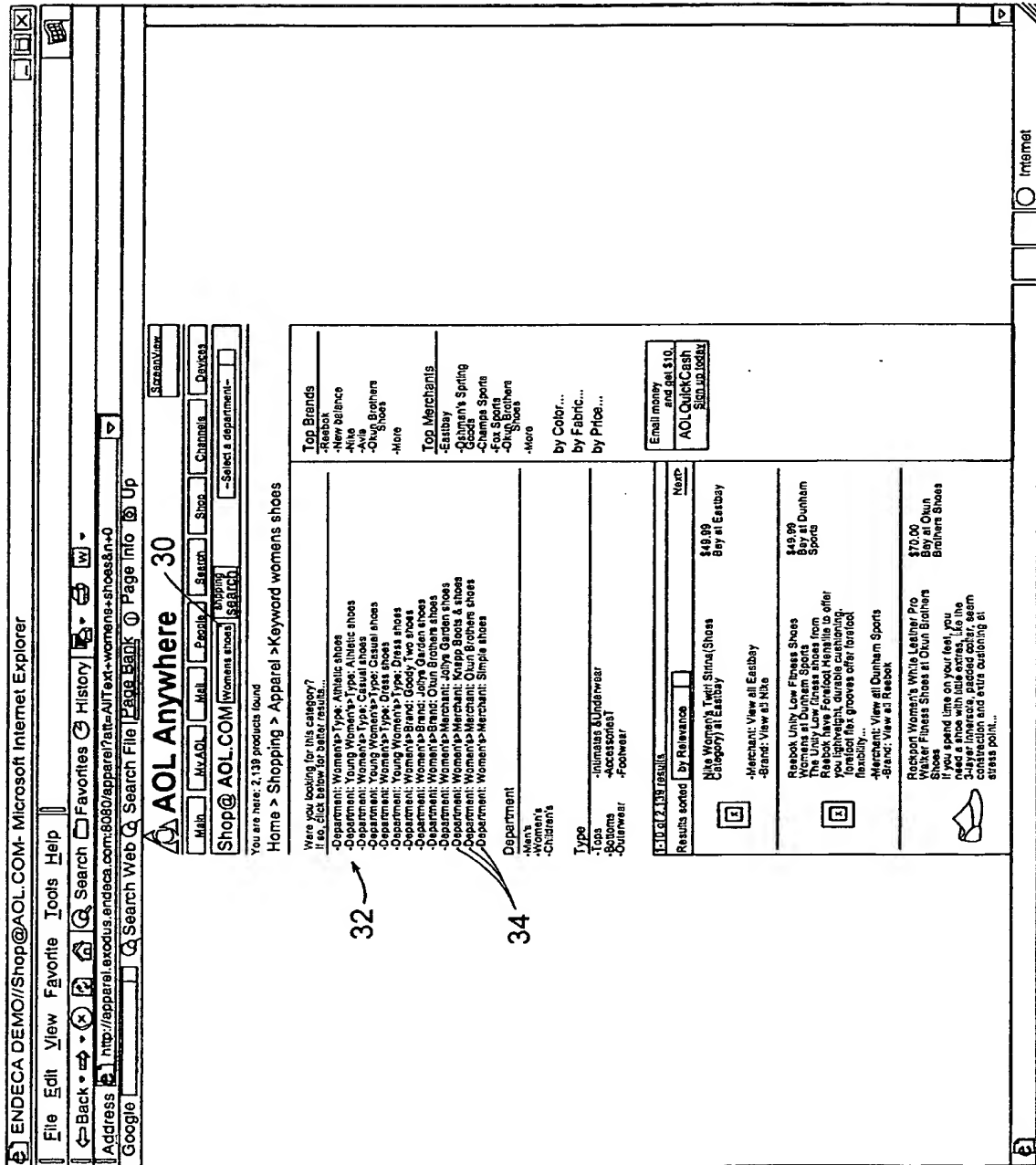
9/25

Fig. 9



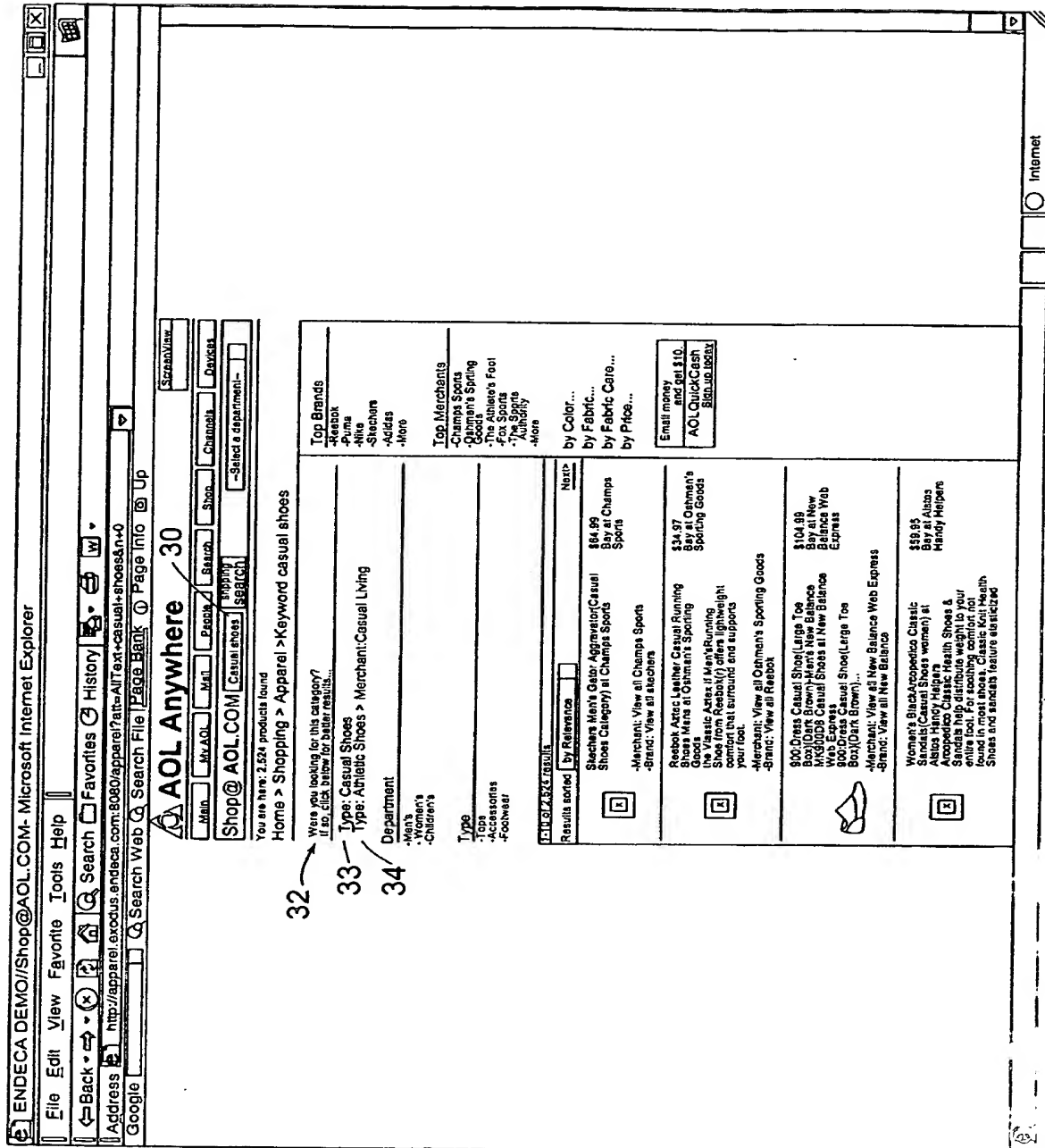
10/25

FIG. 10



11/25

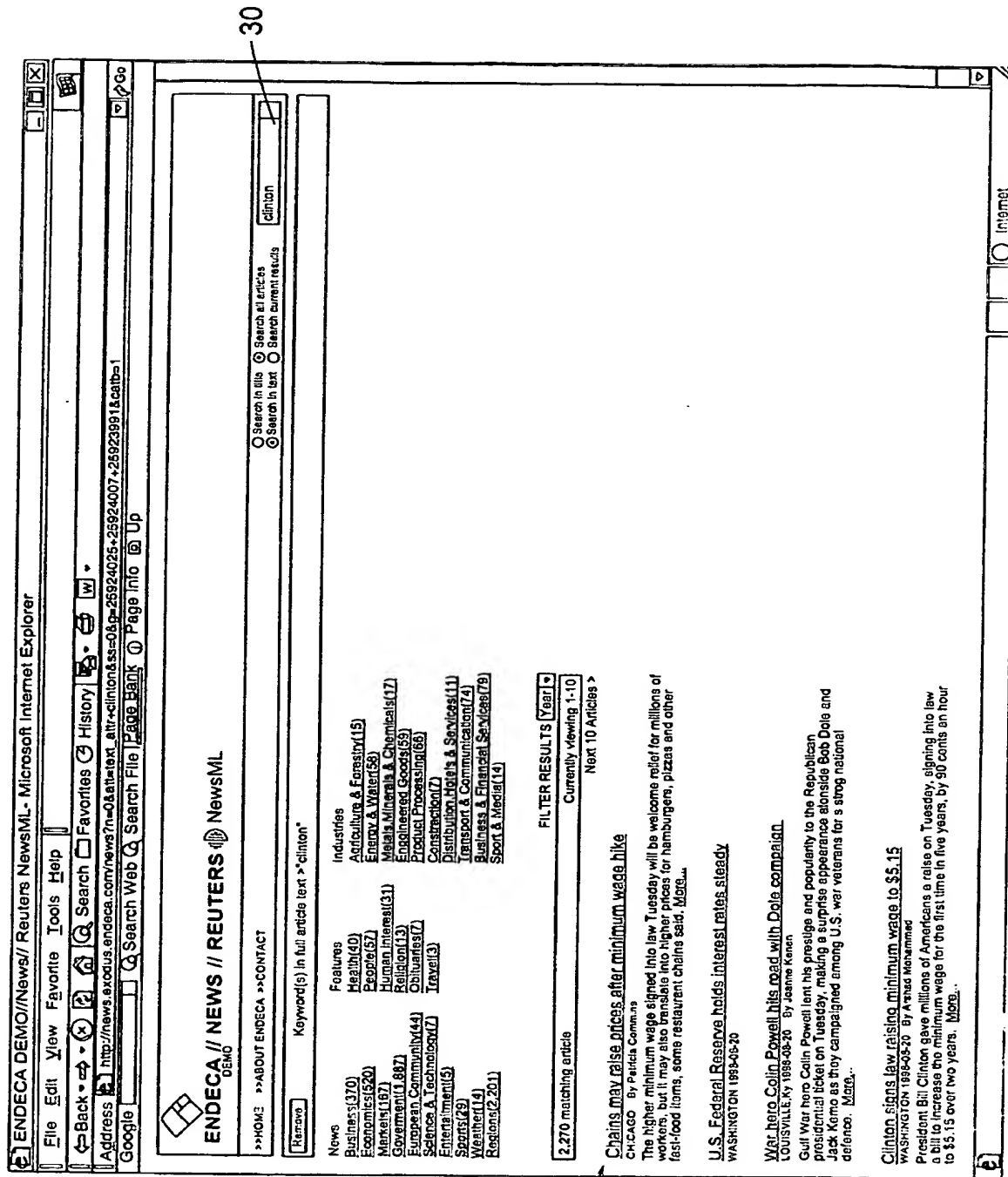
FIG. 11



105

12/25

FIG. 12



35

120

13/25

opti-wine.com

Keyword Search

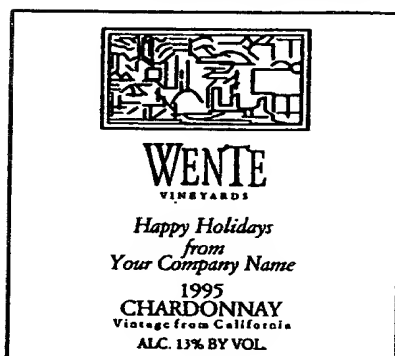
Find

Return to Browsing

30

Chardonnay Monterey County

70



A bold, ripe and full-bodied white from California that offers lots of rich pear, spice, honey flavors, all presented with a light shading of hazelnut. This has a sense of elegance and grace that goes on through the finish. (22000 cases produced)

72

- Wine Types ☐ Chardonnay — 74
 Wineries ☐ Bernardus — 74
 Year ☐ 1994
 Flavors ☐ Hazelnut , ☐ Spice
 Price Range ☐ \$10-\$15
 Regions ☐ US Regions
 Wine Spectator Rating ☐ 90-94

The characteristics above have been used to describe this bottle of wine. Select any combination of different characteristics to see similar bottles of wine...

Build My Store

10

FIG. 13

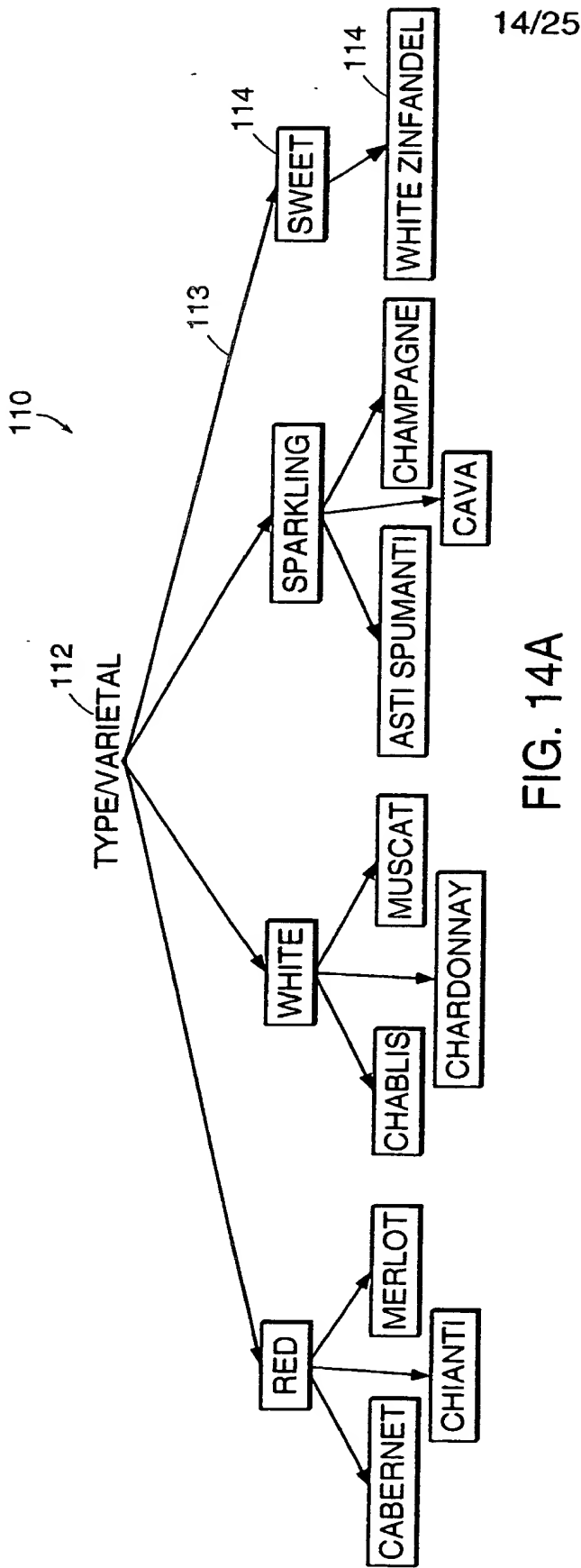


FIG. 14A

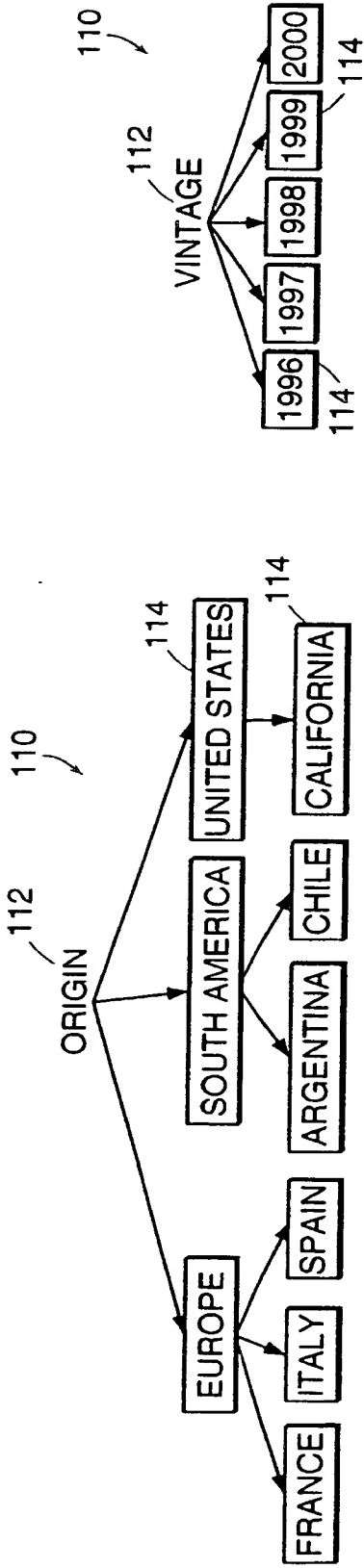


FIG. 14B

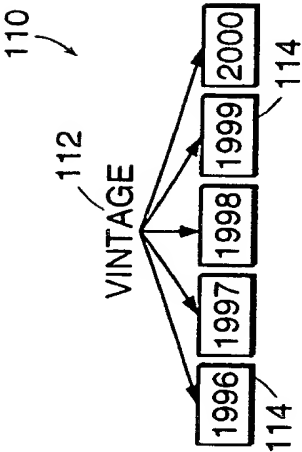


FIG. 14C

15/25

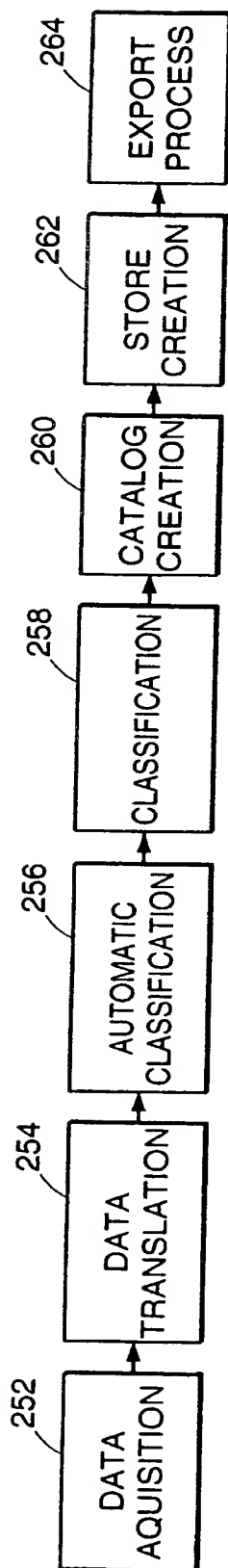


FIG. 15

250

180 → 182 184 186

Doc#	Text Description	Terms
1	Trapiche (award-winning Argentine wine)	Merlot, Argentina, 1998
2	Bully Hill (New York's best)	Chardonnay, United States, 1999
3	Martini & Rossi Asti Spumante	Asti Spumante, Italy
4	Gato Negro ("black cat")	Red, Chile
5	Gato Blanco ("white cat")	White, Chile
6	Chianti "Classico Riserva" Piccini	Chianti, Italy, 1997
7	Gekeikan USA domestic sake	Sweet, California
8	Carnet (Cabernet Franc)	Cabernet, France, 1996
9	Carlo Rossi Chablis (4 L)	Chablis, California
10	Freixenet Cordon Negro (gift box)	Cava, Spain
11	George DuBoeuf Beaujolais Nouveaux	Red, France, 2000

FIG. 16

17/25

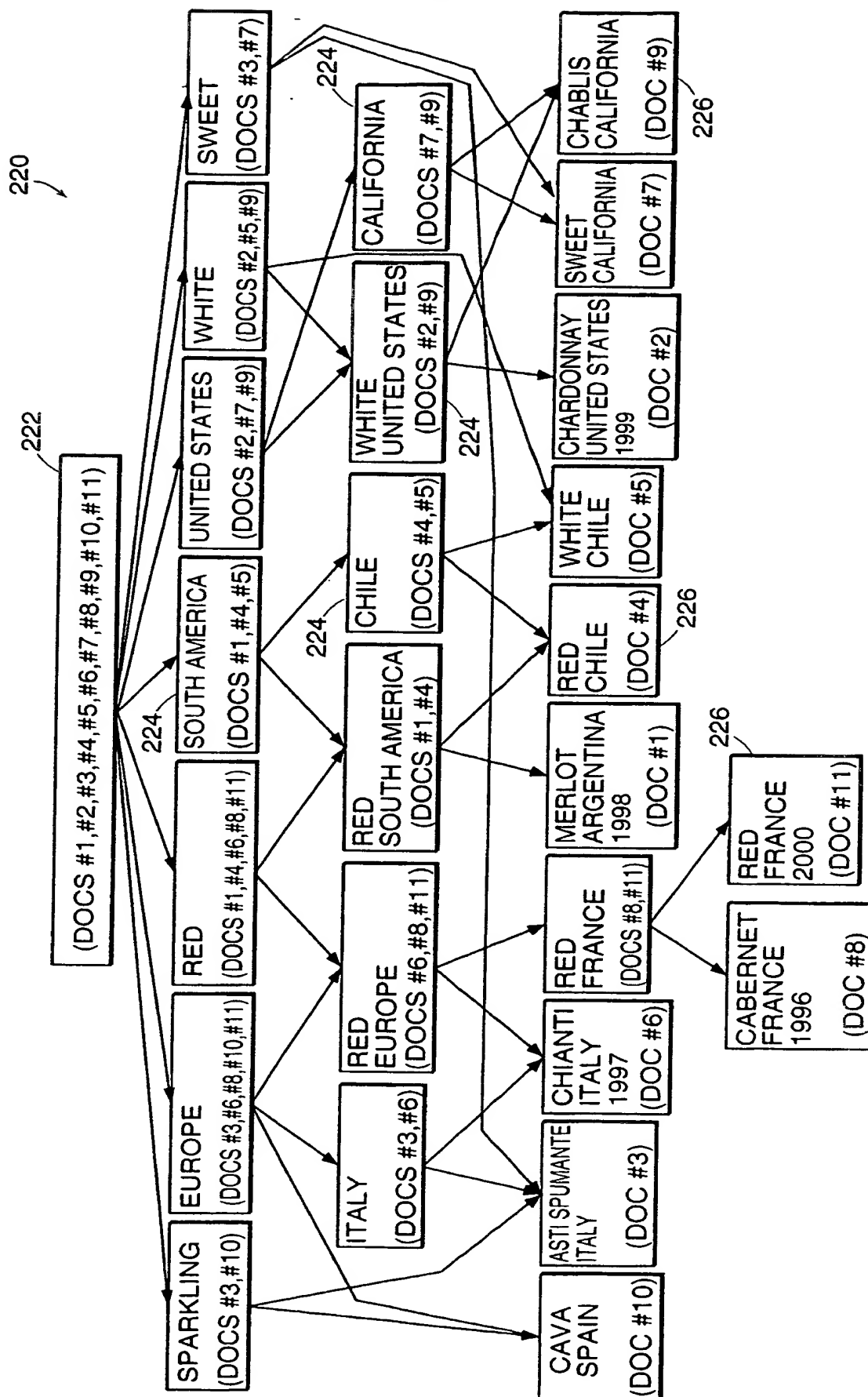


FIG. 17

18/25

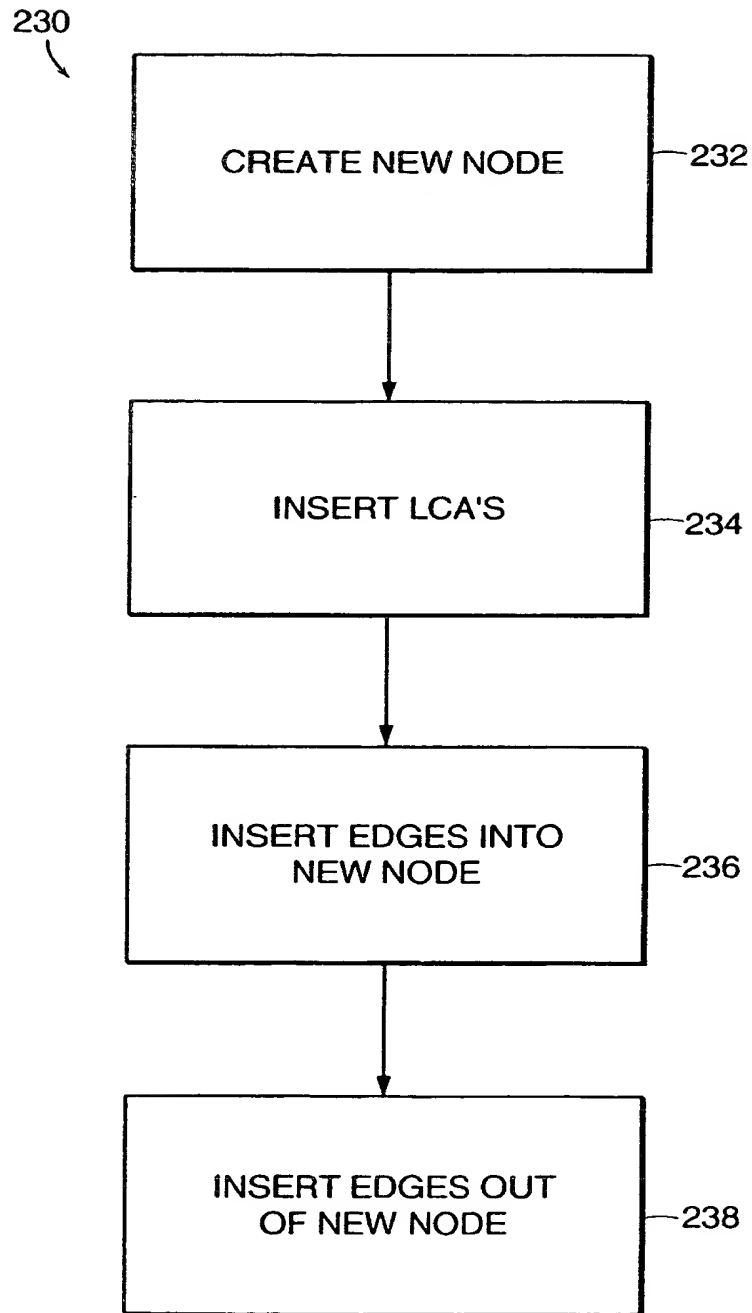


FIG. 18

19/25

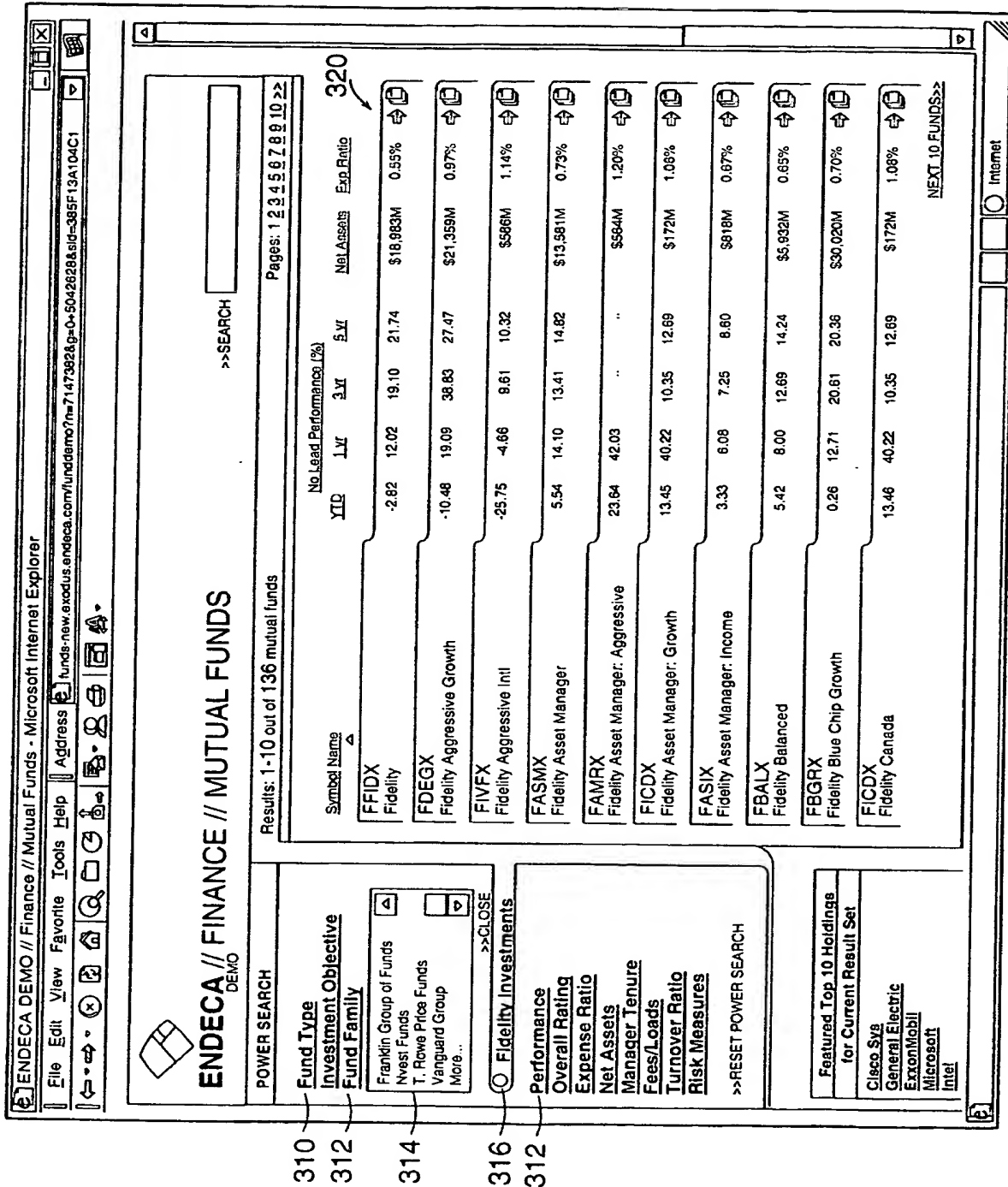


FIG. 19

20/25

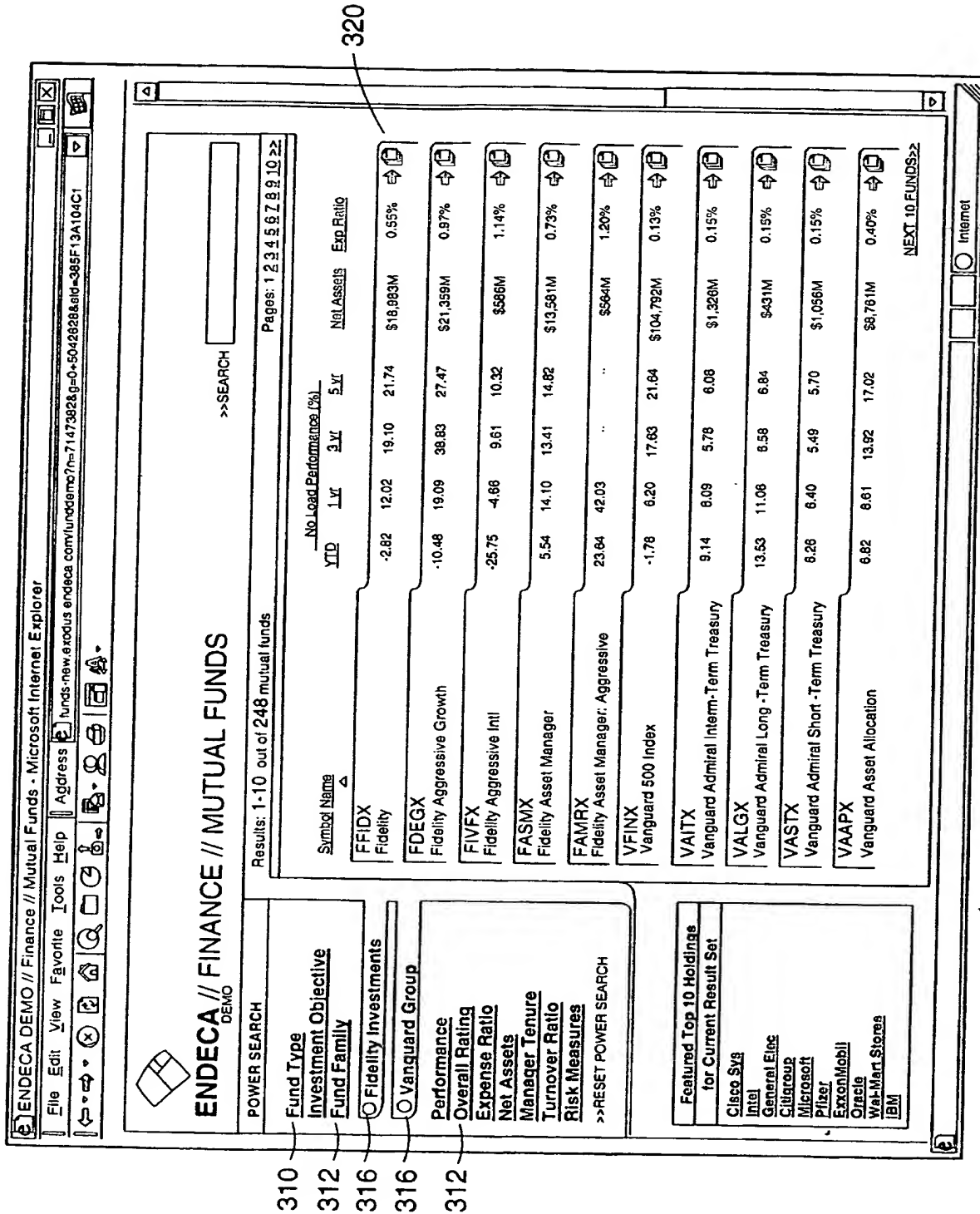


FIG. 20

300

21/25

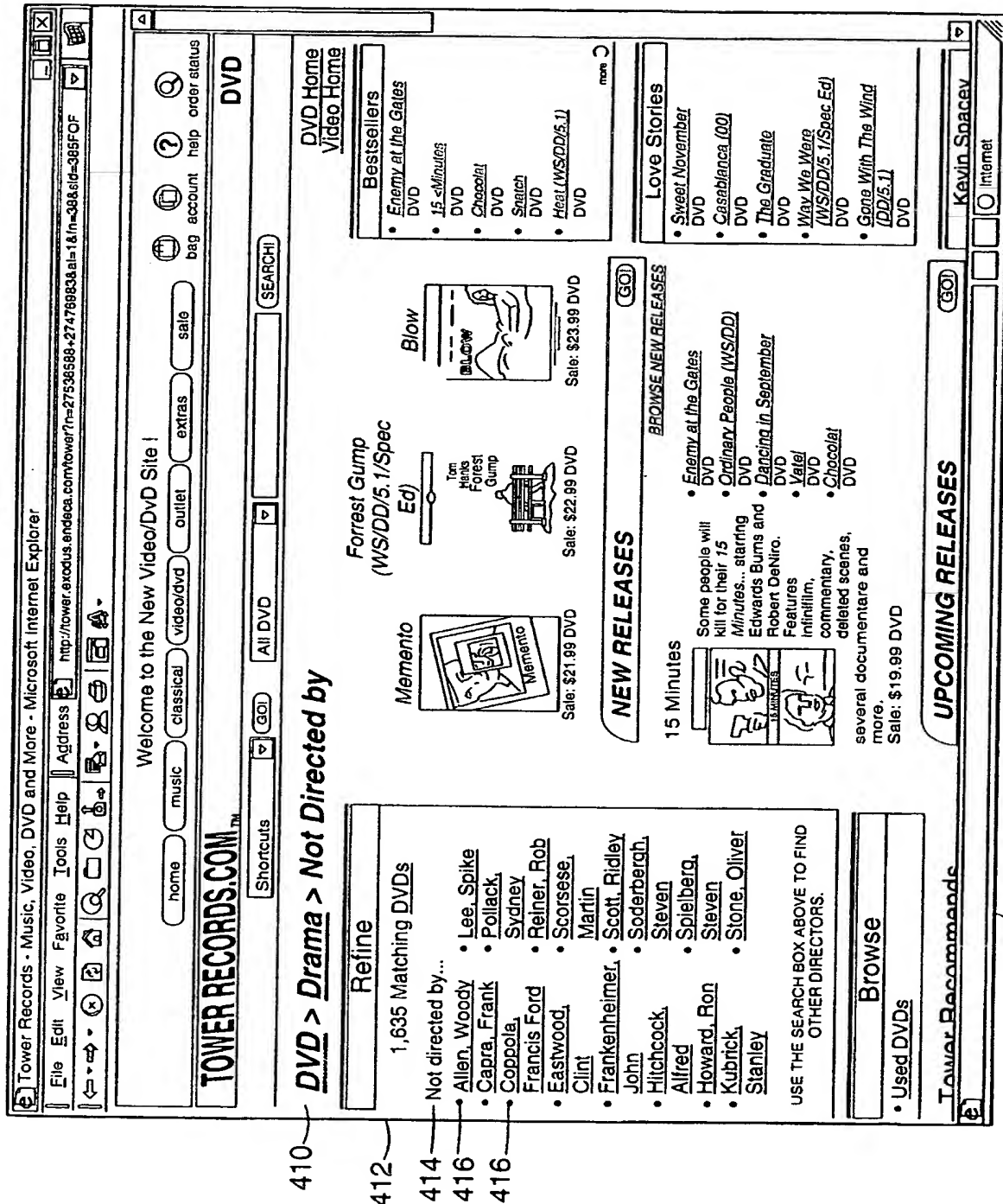


FIG. 21

400

22/25

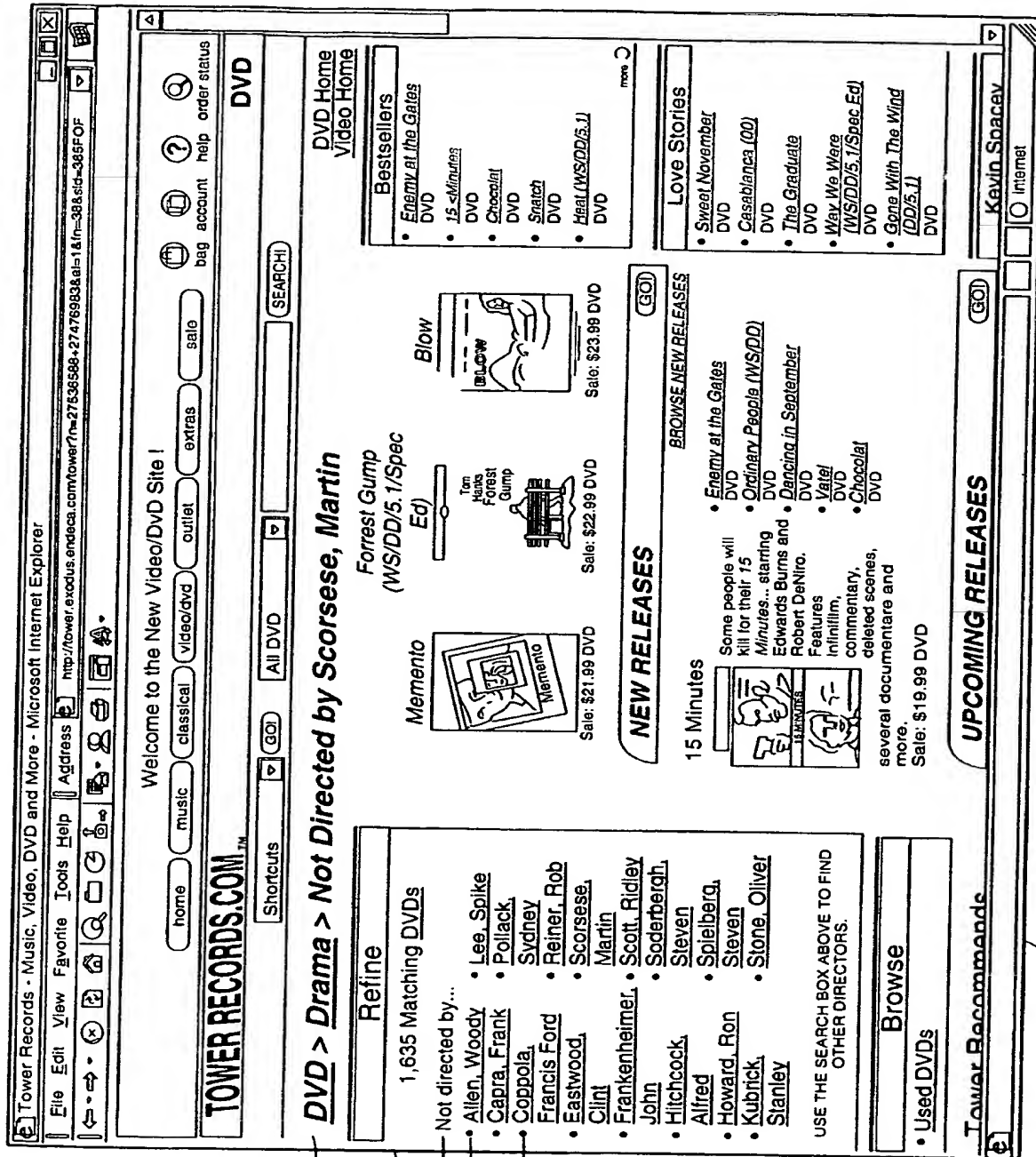
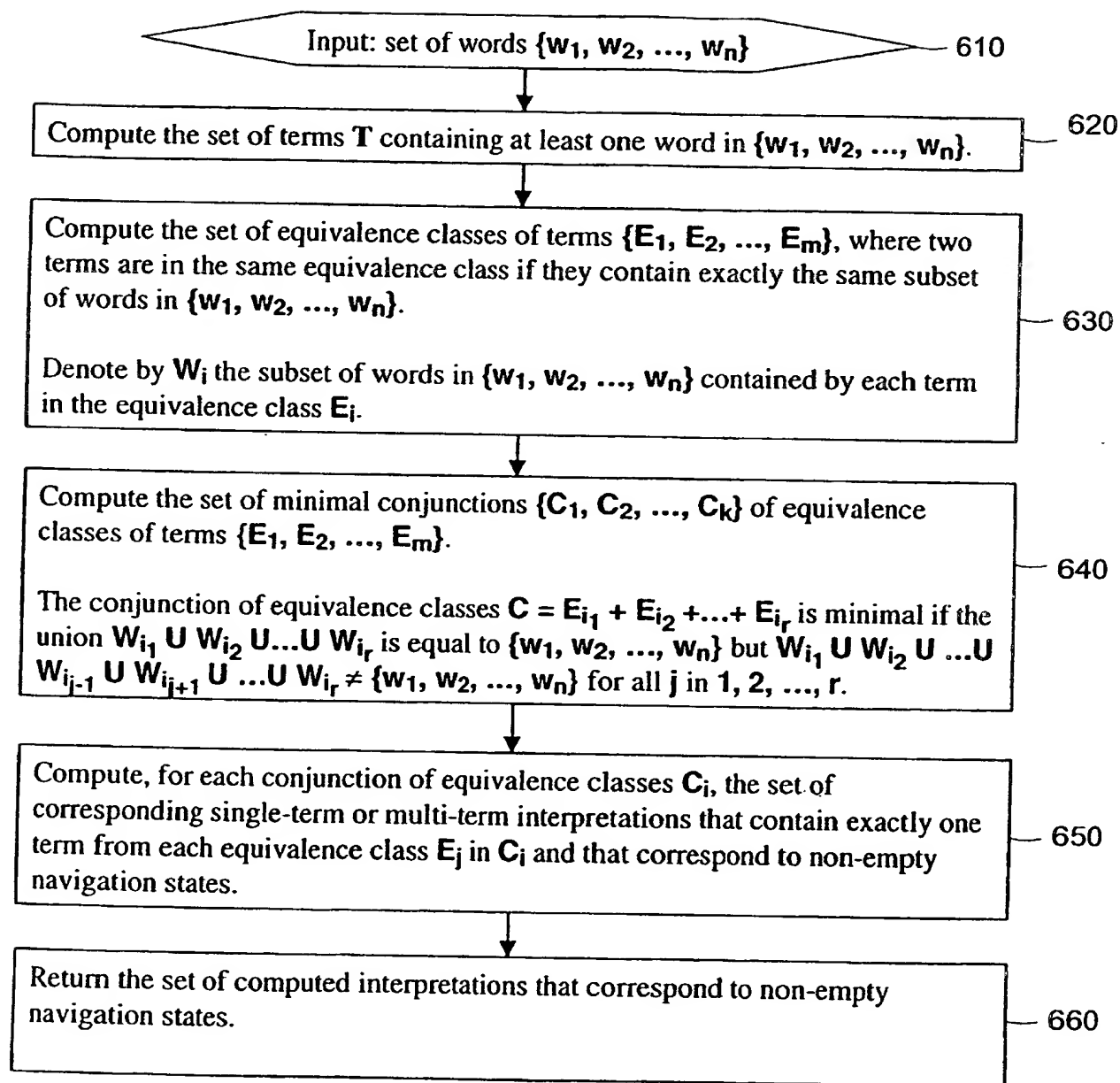


FIG. 22

400

23/25



600

FIG. 23

24/25

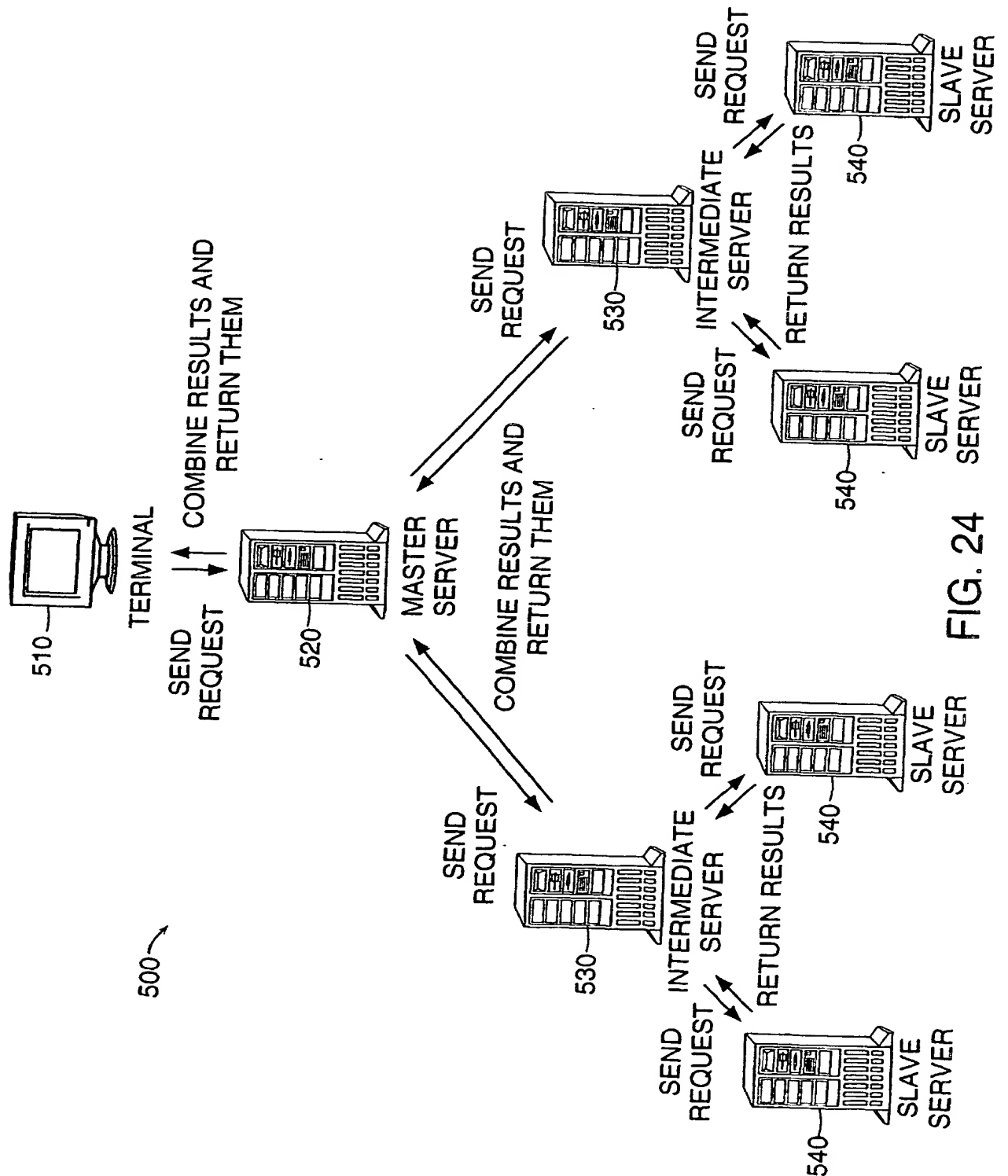


FIG. 24

25/25

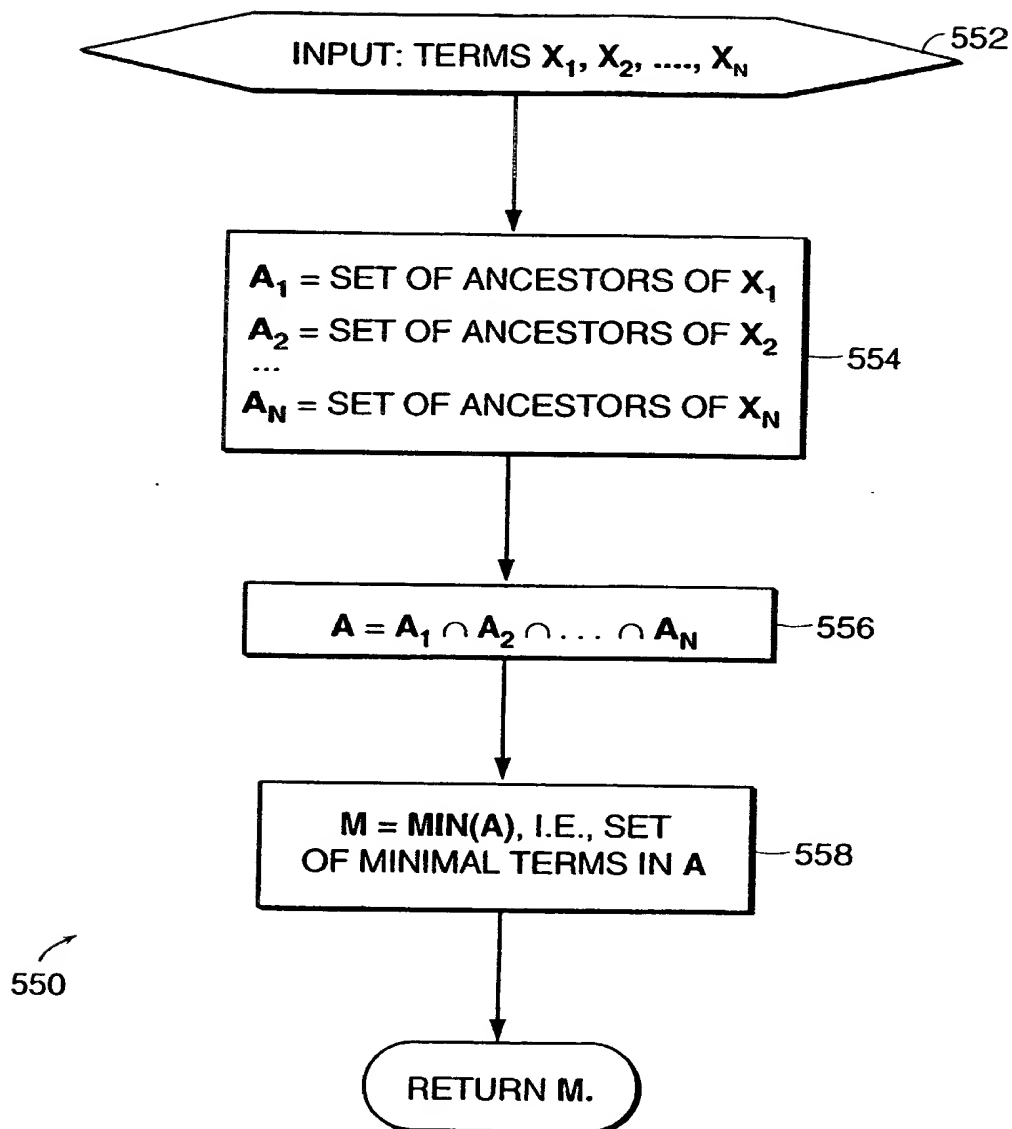


FIG. 25

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 02/25527

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, COMPENDEX, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 01 67300 A (FISH ROBERT) 13 September 2001 (2001-09-13) abstract; figures 1-3A page 3, line 7 -page 4, line 7 page 5, line 15-22 page 6, line 13-23 page 11, line 7 -page 12, line 20 page 20, line 3-16 page 21, line 27 -page 23, line 3 page 26, line 1-7 page 43, line 4-19	1-43
A	US 5 418 948 A (TURTLE HOWARD R) 23 May 1995 (1995-05-23) column 3, line 3-20 column 12, line 55-60	6-8, 23-25, 40, 41



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

8 document member of the same patent family

Date of the actual completion of the international search

15 January 2003

Date of mailing of the international search report

23/01/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Correia Martins, F

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 02/25527

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	POLLITT A S ET AL: "VIEW-BASED SEARCHING SYSTEMS - PROGRESS TOWARDS EFFECTIVE DISINTERMEDIATION" ONLINE INFORMATION. INTERNATIONAL ONLINE INFORMATION MEETING PROCEEDINGS, XX, XX, 3 December 1996 (1996-12-03), pages 433-445, XP008005789 the whole document	1-43
A	US 5 983 220 A (SCHMITT MARTIN) 9 November 1999 (1999-11-09) column 2, line 43 -column 4, line 9 column 2, line 66 -column 3, line 4 column 16, line 16-43	1-43

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 02/25527

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 0167300	A	13-09-2001	WO	0167300 A1		13-09-2001
			AU	3394600 A		17-09-2001
US 5418948	A	23-05-1995	US	5265065 A		23-11-1993
			AU	2801792 A		03-05-1993
			EP	0607340 A1		27-07-1994
			US	5488725 A		30-01-1996
			WO	9307577 A1		15-04-1993
US 5983220	A	09-11-1999	US	6463431 B1		08-10-2002

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.